# RECONSIDERING DISADVANTAGE IN THE UNITED STATES

## AN APPLICATION OF SOCIAL EXCLUSION TO 'BIG' AMERICAN DATA

Dominique L. Green

Doctor of Philosophy in Social Policy

The University of Edinburgh

2019

*To my father, Joseph E. Green, who said I'd be a* ***doctor*** *one day.*
*Continue to rest in peace.*

# Declaration

I declare that this thesis is my own composition, based on my own work, with acknowledgement of other sources, and has not been submitted for any other degree or professional qualification.

Dominique L. Green

# Acknowledgements

‘Thank you’ is the best prayer that anyone could say. I say that one a lot. Thank you expresses extreme gratitude, humility, understanding. – Alice Walker

I would firstly like to extend sincere thanks to my absolutely amazing supervisors: Dr. Paul Norris and Professor Alison Koslowski. My Ph.D. would not be what it is without your unwavering support – personally and professionally – and your belief in me to undertake this ‘big’ project. I am forever grateful. And to Dr. Autumn Roesch-Marsh and Professor Morag Treanor, I thank you both for the additional support you gave. Your advice and interest in my research, along with the opportunities to work with you both have enhanced my Ph.D. experience. I am pleased to call you friends.

Never did I think - coming from little ol’ Johns Island, South Carolina - that I would be as embraced as I have been in Edinburgh, at the University, and amongst a fantastic group of people.

A special thank you to Melissa Highton, who gave me an amazing opportunity to make real change for women in STEM. Your confidence in my abilities has helped me in more ways than I can even explain. Our (award-winning!) work together has undoubtedly influenced this thesis. I am incredibly thankful.

Dr. Heather L.R. Tierney and Dr David Pitts, I thank you both for supporting my application to start this Ph.D. Heather, you spotted a Ph.D. in me from my first course with you in undergrad. Thank you for the encouragement and support for so long. I feel so blessed to know that you’re an email away.

My time in Scotland would not be complete without those I have met along the way. My cohort buddies: Dr Sarah Weakley (DC sister!), Alex Wright, Dr. Patricio Cuevas-Parra, and Dr. Rob Ralston; our chats about PhD life always reminded me that this was not a solo journey. My Edinburgh family: Dr. Dafni Dima, Dr. Katherin Barg, and Gary Mooney - who give me so much comfort. Life here would definitely be a bit dull without you. An extra special shout out to Angelo (Dr. Ercia), whom I met my very first day in Edinburgh. Thank you for the laughs that undid any grey day and the meals you cooked for me while I was too busy ‘living like a student.’ And to Jonathan - thank you for your constant support and company on lazy days.

To my lovely friends in the States who supported my insane idea to move across the pond alone in order to pursue a Ph.D. Special thanks to Charelle Hogan, Jaclyn Weiner, and Michelle Schaefer (to whom I owe so much credit for telling me I could uproot, start anew, and achieve this wonderful thing).

I’m going to put this out there and thank LIZZO! I cannot help it; it’s ‘**cuz I love you**.’ I absolutely do. ‘**Let ‘em say**’ what they want, but I tell you when I listen to your music, I feel ‘**good as hell**.’

Most of all, I thank my mother, Vernelle Green; my sister, Darcelle Green, and my “daughter-niece,” Leila. Your daily video and phone calls (even at 2 am) kept me laughing and upbeat as I spent these years away from you. Your confidence in me to achieve this considerable task means the world to me and I share this with each of you. Mama, especially you, I owe you more than these few lines but all my heart belongs to you. I do my best because I’m counting on you to count on me (Maya Angelou said that before me, but it fits here).

Colossians 3:17

# Abstract

Poverty and disadvantage in the United States are commonly defined in terms of low income. Via this measure, in 2015, over 40 million people were deemed as poor. This definition and subsequent measurement neglect the multidimensional nature of the phenomena. It has been acknowledged that this reductionist measure is insufficient to capture many dimensions of hardship beyond the economic. However, there have been few attempts at quantifying multidimensional disadvantage in the United States.

The aim of this thesis is to quantify multidimensional disadvantage by applying the concept of social exclusion to 'big' American data, the United States Census Bureau-produced American Community Survey (ACS) Public Use Microdata Sample (PUMS) file for 2015 that contains over 2.3 million sample members. Social exclusion, as a concept, theoretically addresses many of the limitations of the official measure. In particular, it offers a multidimensional conceptualisation of disadvantage. This concept, however, is substantially under-researched in the United States. In order to apply the concept to a context in which it is rarely used, social exclusion is measured and defined based on the theoretically derived framework, the Bristol Social Exclusion Matrix (B-SEM). This framework identifies three interconnected domains of social exclusion: resources, participation and quality of life.

The substantive and methodological objectives of this thesis are threefold: 1) to empirically derive the factors of disadvantage in the United States by applying B-SEM to indicators found within the ACS PUMS, 2) to assess the relationship between sociodemographic variables and the dimension(s) of disadvantage, and 3) to explore state-level variation in disadvantage across the United States. An exploratory factor analysis was used to empirically derive the factors of disadvantage in the United States. The results produced three distinct factors: 'labour force participation,' 'economic security,' and 'marriage as a social resource.' This highlights that

disadvantage in the United States is indeed multidimensional, with income representing one component of one factor. Therefore, a focus on a lack of income is incomplete to fully understand disadvantage in the United States. Six ordinary least squares (OLS) multivariate regression models were used to analyse the relationships between the sociodemographic characteristics, age, race, gender, and citizenship status and intersectional characteristics (the intersection between gender and race). In the non-intersectional models, it was found that these characteristics explain more variation in the 'labour force participation' model, compared to the other two dimensions. In the intersectional model, however, over three times the variation is explained in the 'economic security' model, compared to the other two dimensions. The results highlight that different individuals do experience disadvantage differently, showcasing the importance of recognising and addressing multiple forms of disadvantage. Twelve multilevel models were used to assess if there was variation in the dimensions of disadvantage across the United States, if that variation held controlling for sociodemographic characteristics, and if the relationships between the individual characteristics and the dimensions of disadvantage varied across states. The models demonstrated that there was state-level variation in each dimension of disadvantage across the United States and that variation persisted once individual characteristics were controlled for. In addition, it was found that the effect of gender varies significantly across states for each dimension of disadvantage. These results highlight the importance of context in understanding disadvantage and shed light on an important role the state plays in reducing and preventing disadvantage.

These results have important implications for policies designed to alleviate disadvantage in the United States. In addition to expanding all income-based benefits at least to individuals who are 250% above the federal poverty line, state governments should promote the provision of health care to all members of their respective populations and provide incentives that encourage educational attainment.

# Lay Summary

In 2015, over 40 million people in the United States were considered poor. Poverty status is based on a measure developed in the 1960s and has not been updated in the 60 years since. Because the official measure focuses solely on a low level of income, poverty and subsequently, disadvantage in the United States are associated with the amount of income a family has. This project explores disadvantage in the United States beyond the traditional conceptualisation. The main aim of this thesis is to quantify multidimensional disadvantage in the United States via the concept of social exclusion, which has prominence outside of the United States. To this end, using advanced statistical techniques, I analyse the 2015 American Community Survey Public Use Microdata Sample files, produced by the United States Bureau of the Census. A key finding of this research project highlights that disadvantage in the United States is not limited to a low level of income. The analysis uncovered three dimensions of (dis)advantage: labour force participation, economic security, and marriage as a social resource. Another key finding of the study suggests that individuals of various sociodemographic characteristics (such as age, race, gender, and citizenship status) experience disadvantage differently. For instance, the results highlight that Black women in the United States have more advantage in the 'labour force participation' dimension, compared to other groups of women, but have less advantage in the 'economic security' and 'marriage as a social resource' dimensions. Finally, according to the study, there is variation across states in the individual experience of disadvantage. In each dimension of (dis)advantage, the effect of being female varies across states. The study highlights that income is an important, but not the sole component of a complex social issue.

# Table of Contents

# List of Tables

# List of Figures



# List of Equations



# List of Abbreviations

ACS - American Community Survey

PAF - Principal axis factoring

PUMS - Public Use Microdata Sample

EFA - Exploratory factor analysis

| | | |
|---|---|---|
| ICC | - | Intraclass correlation coefficient |
| LL | - | Log-likelihood |
| MAF | - | Master address file |
| MLM | - | Multilevel model/modelling |
| OLS | - | Ordinary Least Squares |
| UK | - | United Kingdom |
| U.S. | - | United States |

# Chapter 1

# Introduction

> Poverty is a "complex set of instances of social exclusion that stretches out over numerous areas of individual and collective existence, and which results in the poor being separated from the generally accepted living patterns in society and being unable to bridge this gap on their own" (Vranken, 2001, p. 75)

Whether it is seen as a manifestation of inequality in the distribution of income (Hill, 1985) or as a distinct phenomenon (Wright and Rogers, 2010), poverty remains a critical social concern for industrial, high-income nations. The way poverty is conceptualised and measured has fundamental implications for how we determine who is poor and why they are poor, but also for how we think about, design, and implement antipoverty policy.

An individual is considered poor in the United States if their family's pre-tax cash income falls below a pre-defined level of income. Via the official poverty measure developed in the mid-1960s by Mollie Orshansky (1965a), there were 43 million Americans classified as poor in 2015 (Proctor *et al.*, 2016). The official measure is frequently "acknowledged to be inadequate for measuring poverty" (Betson and Warlick, 1998, p. 351) for several reasons. For instance, consumption data from the 1950s informs the poverty measure used to determine who qualifies as poor (Fisher, 1997). The measure does not reflect current living patterns or costs of basic needs (Ploeg and Citro, 2008). Additionally, there is a mismatch between the recognised multidimensional nature of poverty and the reductionist measure (level of income) used by the United States to track it (Haveman and Mullikin, 1999). Subsequently, the traditional measure does not capture dimensions of hardship beyond the economic. The unidimensional definition of poverty gives explicit emphasis to the role of income but does not preclude a role for other factors in creating the conditions for social disadvantage more generally. Subsequently, there remains a substantial gap in the literature relating to multidimensional forms of disadvantage in the United States, because research studies often use income to define and measure poverty (e.g. Brady *et al.*, 2013; Cunradi *et al.*, 2000).

Since the development of its official measure of poverty, the United States has not changed its approach to conceptualising and measuring poverty (Glennerster, 2002). However, an important development in the study of poverty has been the shift of emphasis from a single dimension, income, to multidimensional frameworks that encompasses economic and social dimensions (Bossert *et al*., 2012; Engberg-Pedersen *et al*., 2010). For instance, researchers outside of the United States have adopted a view of poverty relative to rising living standards and developed a framework for thinking about nonmonetary aspects of disadvantage (Silver and Miller, 2003). These ideas are embodied in a concept called social exclusion. Social exclusion is used as a framework to conceptualise multiple forms of disadvantage and understand the mechanisms that produce and reproduce it (Babajanian and Hagen-Zanker, 2012).

## 1.1 Thesis overview

The overarching aim of this thesis is to use the concept of social exclusion to quantify multidimensional disadvantage in the United States. Though there have been a few attempts at addressing multiple forms of disadvantage or deprivation in the United States (Dhongde and Haveman, 2016; Wagle, 2009), to the best of my knowledge, none have applied social exclusion to the American[1] context. The literature is limited to a recognition that social exclusion should be examined in the United States (e.g. Brady, 2003; Johnson and Mason, 2012; Silver and Miller, 2003).

[1] I acknowledge here at the beginning that when I state America, I refer to the United States of America. I recognise that this term could be used to represent South America or North America as a whole. However, for the purposes of this thesis, America refers to the United States of America. The term American is then a characteristic of the United States.

Its application here in order to examine any insights into multidimensional disadvantage in the United States is a key contribution of this thesis.

In exploring multidimensional disadvantage in this thesis and gathering an understanding of the factors influencing individual social and economic disadvantage, I seek to answer three main research questions:

1. What is/are the factor(s) of multidimensional disadvantage in the United States?
2. To what extent are sociodemographic characteristics associated with multidimensional disadvantage in the United States?
3. Is there variation in multidimensional disadvantage, on average, across the United States?
    a. Does that variation, if any, still persist after controlling for individual characteristics?
    b. Does the relationship between individual sociodemographic characteristics and multidimensional disadvantage vary significantly across the United States?

### 1.1.1 Detangling terminology (poverty = disadvantage)

In any piece of work, there can be various meanings of words. Poverty is no exception. For the purposes of providing a clear understanding of what is under investigation in this thesis, I offer the following section to untangle what is meant by income poverty, poverty, and disadvantage.

Consider that the United States defines poverty as a low level of income compared to a predefined poverty threshold (background detail is provided in section 1.3). This understanding of poverty is what Chambers (1995) calls 'income poverty,' because it refers to the conventional statistical measures based on income. Income poverty, however, is but one component of disadvantage (Lilja and Watts, 2010). As Vranken (2001) notes in the quote highlighted at the beginning of this chapter, poverty is many, multiple instances of social exclusion. Poverty as social exclusion encompasses various forms of disadvantage, including economic and social ones (Barnes, 2005; Levitas *et al*., 2007). From this perspective, I refer to income poverty as a low level of income. Poverty and disadvantage then are conceptually the same, encompassing the multiple forms of social exclusions an individual may face. This is the focus of this thesis and consequently, these terms are used interchangeably.

I advocate pluralism in understanding disadvantage. Pluralism embraces the diversity of interpretation (Yumatle, 2015). In this thesis, I make no assumption that there is only one approach to understanding poverty (or disadvantage). Social exclusion is chosen because it has been recognised to incorporate much of the most influential conceptual and methodological contributions of the theoretical research on disadvantage (Barnes, 2005). It is chosen to explore if there are any additional insights gained from using the concept within a context in which it is rarely used.

## 1.2 Thesis structure

In addition to the introduction and conclusion chapters, this thesis is structured into three parts or sections: 1) literature and analytical framework, 2) data and methods, and 3) the results. The first section is comprised of Chapters 2 and 3. The next chapter (Chapter 2) critically discusses the literature that has informed my research. Firstly, I acknowledge the limitations of the income poverty measure used in the United States. I then offer social exclusion as a concept that is able to address many of the limitations of the official poverty measure. Following this, I explore why the differences at the individual and contextual levels necessitate a need to abandon a one-size fits all, universal approach to understanding disadvantage, as is currently used in the United States. Chapter 3 presents the analytical framework that encompasses many key characteristics of social exclusion, as highlighted and addressed in Chapter 2. In order to measure social exclusion in this thesis, I have selected to utilise the Bristol Social Exclusion Matrix (B-SEM) framework, developed by Levitas *et al*. (2007). The B-SEM interprets social exclusion as operating within three interconnected domains:

resources, participation, and quality of life. I acknowledge that this taxonomy is not definitive, but it was developed based on an extensive review of the social exclusion literature. As far as I am aware, the B-SEM has not been applied to American data; this is an additional contribution of the thesis. It has, however, been utilised in studies based on data from the United Kingdom (Cusworth *et al*., 2009; Oroyemi *et al*., 2010). I enhance the B-SEM framework by incorporating the concept of intersectionality. Intersectionality, like social exclusion, considers the intertwined economic, social, cultural and political contexts in which individuals and external conditions interact (Cho *et al*., 2013; Crenshaw, 1991; Saatcioglu and Corus, 2014). Informing social exclusion with intersectionality allows me to explicitly recognise how structural characteristics influence disadvantage and how the intersection between characteristics at the individual level can lead to varied experiences of social exclusion.

The second part is comprised of chapters 4 to 6. Chapter 4 presents the data source used for this study. I use the American Community Survey (ACS) Public Use Microdata Sample (PUMS) file for 2015. In this chapter, I recognise the strengths and limitations of using 'big' survey data to measure and analyse social exclusion. I conclude that using this data offers a reliable source to measure social exclusion and provides the opportunity to produce results that can be generalised to the American population. Chapter 5 operationalises social exclusion. In this chapter, I select the indicators from the ACS PUMS that can measure social exclusion as recognised by the B-SEM. Additionally, I recognise the limitations of using a data set that was not meant to measure social exclusion, as some subdomains of the B-SEM do not have an adequate indicator in the 2015 ACS PUMS data. Chapter 6 is the methodology chapter. There, I explain and justify each method employed to answer the research questions acknowledged in section 1.1. I select three methodological techniques 1) exploratory factor analysis, 2) ordinary least squares (OLS) multivariate linear regression, and 3) multilevel modelling to address the substantive and methodological objectives of this thesis.

The final section is comprised of three results chapters. Chapter 7 presents the results of the exploratory factor analysis that uncovers the dimensions of disadvantage in the United States.

Chapter 8 presents the results of OLS multivariate regression models. In this chapter, I examine the relationships between disadvantage and various social categories. Typically, poverty studies within the United States have focused attention on the effects of poverty across social categories that can be generally defined by either race, gender, or age (Norris *et al*., 2010). I recognise that the focus on singular categories neglects a missing piece of the intersection between multiple social identities (Bowleg, 2012; Crenshaw, 1991). Therefore, I apply the concept of intersectionality, as articulated by Crenshaw (1991), in order to explore how identities coexist at the individual level and influence one's experience of various forms of disadvantage. Chapter 9 analyses contextual heterogeneity in the experience of disadvantage. In this chapter, I present the results of multilevel models that examine the variation between states in each dimension of disadvantage. This chapter also explores if the relationship between various sociodemographic characteristics and the dimensions of disadvantage is consistent or varies across the United States.

The final chapter (Chapter 10) concludes the thesis. There, I discuss the contributions and limitations of the research. Additionally, I offer a recap of the findings and provide recommendations on the extension of this research.

## 1.3 Background: The United States measure of income poverty

The United States poverty measure is one of income poverty. This section presents some background on the development of the United States' poverty measure. I offer it here in order to provide a starting place to understand where this thesis seeks to engage with the relevant literature in Chapter 2.

There are two versions of the United States' official poverty measure: 1) the poverty thresholds and 2) the poverty guidelines. The most significant difference between these measures relates to their particular purpose and reporting agency (Fisher, 1992). The poverty thresholds - the main version of the federal measure - are managed by the United States Bureau of the Census (more commonly known as the Census Bureau). They are used primarily for statistical purposes in which thresholds are used to determine who amongst the American population is income poor (Fisher, 1997). The poverty thresholds allow the Census Bureau to classify the poor by a number of characteristics, including age, race, and gender. The poverty guidelines, on the other hand, are handled by the United States' Department of Health and Human Services. The guidelines are used for administrative purposes, in which they are used to decide eligibility for some assistance or service from federally run programs (Fisher, 1997). As the poverty thresholds are the primary measure of disadvantage in the United States, the discussion on the guidelines ends here.

The poverty thresholds were developed in the 1960s by Mollie Orshansky, who worked as an economist at the Social Security Administration. She sought to develop a measure that could assess the risks of low economic status for families with children and to explore how those risks vary among different demographic groups (Orshansky, 1965b). Orshansky advocated that poverty consisted of many different facets, but recognised that there was "…no generally accepted standard of adequacy for essentials of living except for food" (Orshansky, 1965b, p. 5). Therefore, she utilised the United States Department of Agriculture's food plans, which were made for various cost levels. The economy food plan, the cheapest available, serves as the basis of the measure today (Notten and Neubourg, 2007). From there, the development of the measure involved three steps (Orshansky, 1965b), two of which are recognised here.[2]

Firstly, Orshansky (1965b) drafted examples of various family sizes and compositions in order to compute food costs. As we would expect that income requirements would increase as the number of people within the family increased, she made estimates for families that varied in size from two to seven and more. She then classified those estimates based on the gender of the head of the

---

[2] The third step involved comparing cash requirements between farm and nonfarm families. This differentiation between farm and nonfarm household was not a specification between rural and urban living. Even though nearly 20% of the U.S. population lived in rural areas in 2015 (Census Bureau, 2016a), the nonfarm thresholds are applied to rural and urban areas. Subsequently, I keep the discussion about the development of the thresholds to nonfarm households.

household and the number of related children within the specific household under the age of 18. Two-person households were also classified by age in order to take into account older age head of households.

Secondly, she determined the proportion of income assumed to be spent on food and non-food items (Fisher, 1997). Using data that were ten years of old at the time from the United States Department of Agriculture's 1955 Household Food Consumption Survey, Orshansky (1965b) showed that the average American family - households with at least three people - spent a third of its after-tax income on food.[3] This was subsequently applied to situations in which a family would have to cut back on total expenditures in emergencies (Fisher, 1997). The assumption carried that no matter what a family would spend a third of their income on food. Because of this assumption, it meant that the poverty threshold for a family of a particular composition and size was set at three times the cost of the Department of Agriculture's economy food plan (Fisher, 1997). The factor of three, the inverse of the third a family is assumed to pay on food, became known as the multiplier (Fisher, 1997, 1993). For a family of four - two adults and two children, this method resulted in a threshold of just over $3,000 in 1961 ($23,781 in 2015)[4] (Ploeg and Citro, 2008). The family's resources were defined using the Census' money income definition that defined income as all cash before tax (Fisher, 1997; Ploeg and Citro, 2008). This is interesting given that the multiplier is derived based on after-tax income. The family's pre-tax income is then compared to the respective family's threshold in order to determine if they were considered income poor. If the income falls below the respective threshold, every individual in that family is considered income poor.

Interestingly, Orshansky only intended for this work to be used as a research tool (Orshansky, 1968; Ploeg and Citro, 2008), but the thresholds were adopted officially in 1969 when it was designated as the official statistic to be published regularly by the Census Bureau (Fisher, 1997, 1993;

[3] It is important to note that this finding related to families at all income levels and not just the poor (Fisher, 1997).
[4] I obtained this value utilising an inflation calculator available at: https://www.usinflationcalculator.com/.

Johnson and Mason, 2012; Ploeg and Citro, 2008). Since then, there have been some minimal changes in the methodology. The Census Bureau also makes some annual adjustments based on inflation using the Consumer Price Index[5] that does not vary by region or costs of living expenses.

Perhaps due to the consistent and explicit focus on the lack of income in poverty discussions in the United States, there is a belief that boosting income and economic growth would alleviate poverty. For nearly thirty years after the end of World War II, economic growth did decrease official poverty rates (Haveman and Schwabish, 1999). The decreases seen in poverty rates accompanied structural factors, such as rising wages, reduced unemployment, and lower income inequality. In more recent times, the antipoverty effects of economic growth have declined (Dewilde, 2003; Johnson and Mason, 2012). Given the changes in the economy, in the social conditions that affect the demographic composition of the population, and in the public policies that rich countries have seen since the 1970s, the processes of social change have created new social risks (Dewilde, 2003; Hacker, 2006). The United States, in particular, has been transitioning from an industrial to a post-industrial mode of production since the 1970s, facing structural unemployment and new uncertainties in the labour market and an ageing Baby Boomer population (Hacker, 2007, 2006). Increasingly, the income-poor do not necessarily benefit from economic growth because they are typically excluded from labour markets (Galbraith, 2000). All of these factors combine to suggest that poverty should be associated with dimensions beyond income and economic growth. This thesis is an exploration of the multiple forms of disadvantage in the United States that may extend beyond low levels of income.

[5] The Consumer Price Index measures changes in prices that consumers pay for goods and services and reflects the spending patterns for all urban consumers and urban wage earners and clerical workers (Bureau of Labour Statistics, 2018).

## 1.4 Central findings

Is there more than one dimension of disadvantage in the United States? The answer, for the year 2015 (the year of data used), is yes. Based on the findings discussed in Chapter 7, I will argue that there are three dimensions of disadvantage: 'labour force participation,' 'economic security,' and 'marriage as a social resource.' Individual income, which serves as the basis of poverty measurement in the United States, represents but one part of a complex social issue.

In addition to highlighting that there are multiple dimensions of disadvantage for the United States, this thesis will show that different forms of disadvantage are experienced differently by different groups across the United States. Chapter 8 will highlight that women, for instance, are less advantaged than men in every dimension uncovered in this analysis. I will show the importance of adopting an intersectional analysis by offering evidence that the experience of disadvantage for minority women are varied. Black women, for example, in 2015 were more advantaged in the 'labour force participation factor,' compared to White women, but less advantaged in the 'economic security' and 'marriage as a social resource' dimensions. Additionally, I will show that there is contextual variation in multidimensional disadvantage across the United States. In Chapter 9, I will present findings that suggest that even after controlling for individual characteristics, there remains some variation across the United States in multidimensional disadvantage. This evidence will allow me to argue that context (place) plays an essential role in understanding multidimensional disadvantage.

Overall, this thesis will argue that income is not sufficient to capture the multifaceted nature of disadvantage in the United States. Whilst acknowledging Lister's (2004) warning not to downplay income when describing poverty, I will recognise the role that income does play in understanding multidimensional disadvantage. The results of the analysis undertaken in this thesis, however, will offer evidence that the sole focus on income is not complete. I will argue that disadvantage, or poverty broadly is not a unidimensional phenomenon that should be addressed solely through money (such as a basic income) or an increase in the labour force participation rate. It should be addressed in tandem with other issues of disadvantage highlighted throughout this thesis.

# PART 1: LITERATURE AND ANALYTICAL FRAMEWORK

There are two main objectives of Part 1, which is comprised of two chapters. Firstly, I explore the academic literature from which this thesis draws and to which it contributes. Secondly, by engaging with the social exclusion and poverty literature, I discuss the framework used to examine social exclusion in this thesis.

I begin in Chapter 2. Here, I set the stage for my analysis. I offer the reasoning for advocating for an updated conceptualisation and measurement of poverty in the United States. Following on from the background information on the American poverty measure offered in section 1.3, I explore the literature that has critiqued the official measure and the various reasons why it has been deemed inadequate. In doing so, I lay the academic foundations for proposing the use of social exclusion as a lens to conceptualise multidimensional disadvantage in the United States.

Chapter 3 presents, as informed by the social exclusion literature, the analytical framework that guides the measurement of social exclusion. I discuss the definition and measurement of social exclusion via the application of the Bristol Social Exclusion Matrix (B-SEM) to an American context. Additionally, I explore the strengths and weaknesses of utilising the B-SEM framework. I explore how the analysis of social exclusion in this thesis can be enhanced by incorporating the concept of intersectionality to my framework. In doing so, I provide a grounding for later parts of this thesis, including the selection of American indicators based on the B-SEM framework (Chapter 5) and the selection of an appropriate method to examine social exclusion (Chapter 6).

# Chapter 2

# Setting the stage: The American poverty measure needs an update

## 2.1 Introduction

Disadvantage in the United States is conceptualised and measured unidimensionally, discussed as a lack of a sufficient level of income. The purposes of this chapter are two-fold. Firstly, it addresses the reasons why the United States is in need of a reconsidered understanding of disadvantage. In section 2.2, I address the main critiques of the official poverty measure, highlighting why the measure is not fully capable of capturing the multidimensional nature of disadvantage. In addressing these fundamental limitations, I offer social exclusion as an alternative lens to understand disadvantage. Therefore, the second purpose of this chapter is to locate the gaps within the literature this thesis seeks to fill. Section 2.3 introduces the concept of social exclusion, exploring its origins, definitions, and causes. There, I argue that social exclusion theoretically provides a better lens to explore and examine multidimensional disadvantage in the United States. Sections 2.4 and 2.5 explore more practical reasons disadvantage should be conceptualised multidimensionally. Section 2.4 explores how an individual can experience various forms of disadvantage based on their group identity. Section 2.5 acknowledges how where you live can influence individual disadvantage. Section 2.6 concludes the chapter.

## 2.2 The limitations of the United States' measure of poverty

The poverty measure developed by Orshansky (1965b) and addressed in section 1.3 has received significant criticism from much of the American poverty literature (Citro and Michael, 1995; Fox *et al.*, 2015; Fremstad, 2008; Ploeg and Citro, 2008; Segal and Peck, 2006). This section will discuss the major criticisms which centre the measure's focus on food needs, the use of money income to define resources, and a lack of consideration in geographical differences in living costs. Some of these critiques were recognised by Orshansky (1968) herself, but others reflect the changes - social and economic - that have transpired in the United States since the development of the measure in the 1960s.

A consistent critique of the measure is its conceptualisation of need (Ploeg and Citro, 2008). Doyal and Gough (1991) recognise that there is more than one type of human need. They claim that physical survival and personal autonomy are the most basic of human needs because they represent the conditions for any individual action (Doyal and Gough, 1991). The United States' measure focuses solely on one type of need: food need, which in some respects is related to physical survival. The basic critique of the American measure is the sole emphasis on food need. Other needs that were important in the 1960s are not considered such as health care, clothing, or housing (Ploeg and Citro, 2008). Need, in the decades since the development of the measure, now includes items, such as indoor plumbing and telephones. In addition, other expenses are now noted to take up more proportions of total income than they did 50 years ago. Housing, for instance, makes up a much larger share of household budgets (Hutto *et al*., 2011). In 2015, housing expenditures in the United States ranged from 31 to 37% of total income, whereas food expenditures ranged from 12 to 13% (Bureau of Labour Statistics, 2016). This reflects Haber's (1966) suggestion that the amount of family income spent on food was overestimated in the 1960s. Orshansky (1968, 1965b) did recognise that this was a limitation of the measure and suggested continuous updates of the thresholds to reflect these types of changes. Her suggestion, however, has not been implemented.

Another critique of the official measure is that solely money income is used to define family resources. The measure only includes money income and transfers based on income before taxes (Ploeg and Citro, 2008). Money income is noted to be a weak measure of the consumptive power of the income poor since many receive in-kind benefits, such as Medicaid,[6] food benefits, or even emergency assistance from government agencies or charities (Citro and Michael, 1995). In-kind benefits, be it public or private, tend to raise the living standards of the poor (Smeeding, 1977). However, in determining poverty status, these types of benefits are not counted as income (Tiehen and Ploeg, 2012). Consequently, income and subsequently, the poverty measure are not sensitive to changes in these programs (Getz, 1985; Ploeg and Citro, 2008; Smeeding, 1977). This critique offers further evidence of the lack of change in the measure. When Orshansky (1965b) developed the measure, in-kind benefits for families with low incomes came in the form of cash transfers and individuals with low levels of income paid little taxes at the time (Ploeg and Citro, 2008). Over 50 years later, in-kind benefits have grown and a sizable portion of America's anti-poverty strategy comes in the form of the earned income tax credit (Ellwood, 2000).[7] These are not included in the measure.

Because the United States' definition of poverty gives explicit emphases to the role of income, it does not preclude a role for other factors in creating the conditions for social disadvantage more generally (Conley, 2005). There is sufficient argument in the literature against defining and measuring poverty with a single dimension, in terms of monetary income (Betti and Verma, 1999; Haveman and Mullikin, 1999; Townsend, 2006; Vranken, 2001). The literature increasingly recognises that poverty is a complex and multidimensional social problem with various economic and non-economic manifestations (for instance, Adams and Page, 2001; Atkinson, 2003; Blank, 2003; Chakravarty, 2006; Rank, 2006; Smeeding *et al*., 1993). Though using money income can reflect a family's ability to meet its immediate needs, it indicates little about the level of consumer spending that is available to that family (Haveman and Mullikin, 1999). The recognised multidimensionality implies that we can use different indicators besides income as a welfare indicator.

---

[6] Medicaid is a federally assisted, but state administered health financing program for low income individuals in the United States (LeBlanc *et al.*, 2001; Snowden and Thomas, 2000).
[7] The Earned Income Tax Credit is a refundable tax credit for low income workers (Ellwood, 2000; Sykes *et al*., 2015).

Consider that since the 1960s, the United States has gone through meaningful social and economic changes that cannot be (or are not) accounted for with the current measure. For instance, women's labour force participation has risen substantially (Treas, 1987). This has increased the costs of childcare associated with mothers in the labour market and is subsequently likely to be a more substantial component of family budgets (Hofferth and Wissoker, 1992; Ribar, 1992). These changes and other conditions of social disadvantage are not fully reflected in the current official poverty measure of the United States (Ploeg and Citro, 2008), but should have substantial implications for how poverty is conceptualised and measured.

Lastly, the U.S. poverty threshold does not consider regional variation in the cost of living (Berube *et al.*, 2012; Citro and Michael, 1995; Nelson and Short, 2003). Living on a yearly income of $10,000 is substantially different in a city in a southern state like Mississippi than it is in an affluent city, such as Los Angeles. Additionally, the definition of the family unit in counting resources has also been critiqued (Ploeg and Citro, 2008). The measure excludes unmarried cohabiting partners and other members of the household who are not related (Kenney, 2004). Currently, the measure includes thresholds for only related individuals living in the same household.

***The Supplemental Poverty Measure***

The most salient attempt, so far, at addressing limitations of the measure has been via the development of the Supplemental Poverty Measure (SPM). The SPM, however, still relies on money income to determine poverty status and fails to address the non-economic manifestations of disadvantage. Next, I will discuss the SPM, while addressing its own limitations. I acknowledge that the SPM is also not adequate in fully helping us to understand the multifaceted nature of disadvantage.

After several calls for an updated measure (Blank, 2008; Citro and Michael, 1995; Ruggles, 2008), the Obama Administration (2009-2017) approved plans in February 2010 to develop this new experimental version of the poverty thresholds based on Citro and Michael's (1995) report that addressed some criticisms of the measure (Hutto *et al.*, 2011). This experimental measure, called the Supplemental Poverty Measure (SPM), does not replace the current poverty measure but is instead, as the name implies, a supplement.

The SPM and the official measure have notable similarities and differences. Like the official measure, the SPM is a measure of economic disadvantage (Dalaker, 2017). The SPM defines an individual's or family's poverty status by comparing resources against a measure of need. The SPM and the official poverty measure have various poverty thresholds to account for family size and composition. If a family has fewer resources compared to their specific thresholds, they are considered poor. Unlike the official measure that measures poverty in absolute terms, the SPM is 'quasi-relative' (Johnson and Smeeding, 2012). The SPM is absolute because the thresholds represent a dollar amount spent on a basic basket of goods that include clothing, shelter, food, and utilities, and a small sum for personal needs (Short and Garner, 2012). It is relative because it is updated to account for the changes in the costs of basic needs (Citro and Michael, 1995). The additional differences between the two reflect the changes in household composition since the development of the original measure and in their definitions. Family is defined based on an approach that considers how household members share the resources within the household (Dalaker, 2017). In addition, the SPM geographically adjusts need based on housing costs by metropolitan area or by state for non-metropolitan areas and by homeownership to reflect differences in housing costs (Dalaker, 2017).

Though the SPM was designed to address some critical limitations of the official measure, there are two distinct but related difficulties. Firstly, like the official poverty measure, the SPM relies on annual money income as the indicator of resources and ignores many other potential sources of welfare, such as social inclusion that may be less directly tied to income (Haveman and Mullikin, 1999). Secondly, the SPM still determines the poverty status of Americans by comparing their financial resources against poverty thresholds. The SPM, as presented by Citro and Michael (1995, p. 19) interprets poverty as a "lack of economic resources (e.g., money or near-money income) for consumption of economic goods and services (e.g., food, housing, clothing, transportation)." Essentially, the SPM takes a reductionist approach to understand the complexity of poverty just as the official measure does (Wagle, 2002). The SPM does not take into consideration the non-economic manifestations of poverty.

Considering its limitations, the traditional approach to measuring disadvantage has endured. It has important descriptive value, offering a relatively easy way to compare rates across groups and over time. It is an objective measure that can be easily discussed and analysed in policy and academic arenas. However, the intricacies of other disadvantages, such as unmet medical need, food insecurity, and unequal access to educational attainment are all essential aspects that warrant exploration. These cannot be explored using income solely as the measure of disadvantage. In addition, the individual experience of disadvantage can be varied based on the various characteristics of the individual. The United States poverty measure offers a one size fits all approach that is not equipped to consider these variations. This study seeks to tackle this gap in the measurement of disadvantage in the United States by addressing these limitations of the measure. The rest of this chapter addresses these gaps. In the next section, I introduce the concept of social exclusion that has the theoretical underpinnings to expand American notions of disadvantage and addresses the limitations of the measure. Section 2.4 addresses the variation of disadvantage at the individual level. Section 2.5 addresses the importance of analysing contextual geographic differences as has been previously recognised with the development of the SPM.

## 2.3 Social exclusion: a lens to examine the multidimensionality of disadvantage

Discussions of disadvantage - particularly outside of the United States - have turned to more pluralistic approaches and interpretations. For instance, the capability approach has been used to conceptualise multidimensional disadvantage in developing and developed countries (Alkire, 2005; Sen, 1999). Additionally, social exclusion has also gained prominence in the United Kingdom (Crous and Bradshaw, 2017; Main and Bradshaw, 2014; Oroyemi *et al.*, 2010), Australia (Saunders, 2013, 2003), and across other European countries (Pirani, 2011; Vranken, 2001) These pluralist interpretations of poverty recognise that disadvantage is marked by more than a lack of economic resources, but rather is a complex social problem with various economic, social, political, and cultural manifestations (Ravallion, 1996; Saatcioglu and Corus, 2014; Silver and Miller, 2003).

I do not utilise the capability approach in this thesis to understand multidimensional disadvantage, despite its popularity (Agee and Crocker, 2013; Alkire, 2005, 2005; Burchardt and Hick, 2018; Giraud et al., 2013). This choice is made for two reasons. Firstly, two studies exist examining multidimensional capability deprivation for the United States (Dhongde and Haveman, 2016; Wagle, 2009). Secondly, though both social exclusion and capability deprivation are fundamentally concerned with multidimensional disadvantage (Levitas *et al*., 2007), the added benefit of social exclusion is that there is an explicit focus on the relational processes of impoverishment, the structural characteristics that are responsible for disadvantage, and the group issues that are often neglected in other approaches (Laderchi *et al*., 2003; Sen, 2000). The contribution of this thesis is to assess social exclusion, which to the best of my knowledge has been analyzed in few studies for the United States context (for instance, Lee and Cagle, 2018), but not for the time frame under investigation, and to examine any insights gained from this conceptualisation of multidimensional disadvantage. The following section explores the theoretical contributions of social exclusion and makes a case for using social exclusion in the United States as it addresses many of the critiques of the United States measure.

### 2.3.1 What is social exclusion? History and definitions

The emergence and modern conceptualisations of social exclusion have been frequently attributed to the work of Rene Lenoir[8] (Bossert *et al.*, 2007; de Haan, 2000; Levitas, 2006; Mathieson *et al.*, 2008; Silver, 1994; Silver and Miller, 2003). Lenoir used the term '*les exclus'* to refer to individuals who fell through the social insurance system safety net, such as the homeless and lone parents, and those who appeared to be at the margins of French society (Morgan *et al.*, 2007). Social exclusion came to signify a rupture of the social bond between the state and its citizens (Silver and Miller, 2003). This disconnect was extended beyond income poverty to include characteristics such as a lack of participation in politics, geographical isolation, and poor health (Davies, 2005).

Since then, the term has emerged with reference to problems related to a new form of poverty. This idea of new poverty suggested that the issue is not just monetary, but also includes profound historic changes both at individual and societal levels. The new poverty touches on issues of unemployment, the concentration of the disadvantaged in segregated areas, poverty among young adults, amongst many other issues (Baulch, 1996; Chambers, 1995; Schierup, 2010).

Like many concepts, there is no generally agreed-upon definition of social exclusion. Silver, (1994, 2007) suggests that the meaning of social exclusion and its use varies across countries because it is rooted in different traditions and political histories. Subsequently, it can have multiple meanings in different contexts and for a number of different purposes (Brady, 2003; Silver and Miller, 2003). Despite this, there has been some consistency within the literature about the primary elements of social exclusion that are worth exploring. It is multidimensional, relational, dynamic, and relative (Atkinson, 1998; Room, 1995).

Most importantly, there is agreement that social exclusion is a multidimensional concept (Barnes, 2005; Pirani, 2011; Room, 1995; Silver and Miller, 2003; Silver, 2007). This characterisation of social exclusion considers the various ways in which disadvantage operates to limit the

[8] The original work by Lenoir (1974) is written in French. Consequently, I relied on research written in English that highlighted Lenoir's contribution to the concept of social exclusion.

opportunities and life chances of various individuals and groups (Sonowal, 2008). It implies that a person can be excluded from housing, employment, minimum consumption, livelihoods, property, etc. (de Haan, 1999; Silver, 1994). It also means that a person can experience multiple disadvantages at the same time. From this viewpoint, an examination of a person's living standard or quality of life cannot be based on just economic indicators. It points to a need to also examine the influence of social relationships (Pirani, 2011). This can include social support networks and social participation, which can have a positive influence on other areas of an individual's life. Weak social interactions with family and friends and with the broader community represent forms of social disconnectedness that can keep an individual excluded in various other areas of life, including employment (Cornwell and Waite, 2009).

Social exclusion is also relational. It arises as the product of social interactions that can be characterised by unequal power relations (Mathieson *et al.*, 2008). It can produce breaks in the relationships between people and society, which then results in an absence or loss of social participation, social protection, social integration, and also power (Silver, 1994).

Social exclusion is also a dynamic concept. It examines the processes that contribute to exclusion as opposed to income poverty that is characterised as a static outcome (Barnes, 2005). Thought of this way, social exclusion does not arise simply from an individuals or groups current status, but it is connected to their past background and prospects for the future. Additionally, social exclusion is concerned with the institutional rules and relationships that determine the distribution of resources and assign value in society (Kabeer, 2005). It focuses particularly on the mechanisms that grant or deny access and recognition. Thus, it can involve the regular denial of entitlements to resources and services and the denial of the right to participate equally in relationships in economic, social, political, or cultural arenas of the society in which an individual lives (Islam, 2015). In this sense, social exclusion also encompasses agency. It points to who is doing what in relation to whom (Saunders, 2013). Therefore, exclusionary processes can occur at various levels – within and between households, cities, and states (Engberg-Pedersen *et al.*, 2010; Islam, 2015; Kabeer, 2000).

Another important element of social exclusion is its relativity. The concept suggests that a person is only excluded in the society in which they live at a given time (Atkinson, 1998; Bossert *et al.*, 2007). There does not exist an absolute social exclusion condition. This aspect of social exclusion is not a recent articulation in the literature as understandings of disadvantage had been evolving for some time outside of the United States before the term social exclusion appeared in European lexicon. In the late twentieth century, a group of social scientists turned to style of living and the relative deprivation approach to understand poverty. Townsend (1979) introduced the idea of relative deprivation, stating that:

> "Individuals, families, and groups in the population can be said to be in poverty when they lack the resources to obtain the types of diet, participate in the activities and have the living conditions and amenities which are customary, or are at least widely encouraged or approved, in the societies to which they belong. Their resources are so seriously below those commanded by the average individual or family that they are, in effect, excluded from ordinary living patterns, customs and activities" (Townsend, 1979, p. 30).

Townsend (1979) broadened understandings of poverty as something more than just a lack of income. The relative deprivation approach rests on the idea that if people are deprived to such an extent that they lack the resources to participate in society's customary activities, and subsequently, in some sense they are excluded from society, then they are believed to be in poverty (Madden, 2000). The traditional notions of poverty in the United States are based on basic needs, particularly adequate food. However, as Madden (2000) notes, Townsend's (1979) idea of deprivation embodies a relative concept of poverty, because since poverty is defined as exclusion from the norm and we assume that the norm changes over time, then so should the definition. Townsend (1979) did not use the term social exclusion, but he did argue for a more complex analysis of impact of poverty on an individual's life, thereby establishing the relationship between material disadvantages and not being able to participate in society fully. As such, the only way in which we are able to judge if a person is excluded is to observe that individual relative to the context and the society in which she lives. These, for the purposes used in this thesis, are further expanded upon in section 2.4 for individual disadvantage and section 2.5 for contextual variation in disadvantage.

These attributes aim to conceptually differentiate the traditional notion of income poverty from social exclusion, as it is important to stress that income poverty and social exclusion are not just interchangeable concepts (Atkinson and Hills, 1998; Vranken, 2001). Via these characterisations of social exclusion, social exclusion links together both social and economic disadvantages. Therefore, it encompasses not only the lack of access to goods and services which underlie the traditional notions of poverty, but also exclusion from security, from participation, and from representation. It addresses components of disadvantage that makes it more suitable to capture the multifaceted nature of disadvantage in the United States.

### 2.3.2 The causes of social exclusion: various discourses

This section discusses the types of social exclusion as it has been discussed by Levitas (2005) and Silver (1994). Silver (1994) offered three paradigms of social exclusion that acknowledge that the meaning and causes of social exclusion will vary in different contexts. Levitas (2005), on the other hand, identifies three different positionings to social exclusion in British public policy that implies different models of causality and different policy interventions that, as she suggests, embed the causes of social exclusion (Levitas *et al.*, 2007). Both lines of discussion describe the distinct approaches to explaining the various causes of social exclusion. I start with Levitas' (2005) discourses of social exclusion.

The three discourses of social exclusion as per Levitas (2005) are the 1) social integrationist discourse, 2) the moral underclass discourse, and 3) the redistributionist discourse (Levitas, 2005; Levitas *et al*., 2007). Firstly, the social integrationist discourse suggests that paid work and employment are most important for social inclusion (Levitas, 2005). This approach narrows social exclusion to a lack of participation in paid work. A focus, however, exclusively on paid work ignores the multidimensionality of social exclusion. It also ignores the inequalities between people who are in paid work. It also does not recognise the significance of unpaid work, such as childcare that is frequently undertaken by women (Watt and Jacobs, 2000).

The moral underclass discourse is mainly concerned with morality in the behaviour of the excluded themselves. It presents socially excluded individuals as culturally distinct from the mainstream (Levitas, 2005). This discourse essentially merges the concept of social exclusion with the concept of the underclass that has existed in American political and academic discourse since the 1960s (Michels, 2013; Wilson, 1987, 1984). The underclass refers to a group of people at the bottom of the socioeconomic structure (Kronauer, 1998; Wilson, 1987). This group, as per various explanations, developed their own lifestyle that separates them from the rest of the population. Different writers have emphasised culture (Murray, 1996, 1994), structure (Wilson, 1987), or racial segregation (Massey and Denton, 1993) as the main driving forces behind this 'separation' from the mainstream society, but the underclass are frequently blamed for their situation. Because discourse exists that blames the individual for his or her own exclusion, there is potential for the concept of social exclusion to face the same problems as that of the underclass. In order to shift the focus beyond characteristics of those excluded, the literature stresses that social exclusion is indeed an outcome but it is also a process where actors or institutions are responsible for producing exclusion (Kabeer, 2005; Kronauer, 1998; Levitas, 2006; Silver, 2007).

Lastly, the redistributionist discourse asserts that the lack of resources, not just the lack of money, is the key problem and is therefore concerned with poverty (Levitas, 2006). This discourse stresses access to services and lack of full citizenship rights as the leading causes of social exclusion (Levitas, 2005). The redistributionist discourse recognises that people can be excluded due to discrimination on the grounds of their characteristics (Watt and Jacobs, 2000).

Silver (1994) presents three 'paradigms' which are based on different notions of social integration. She identifies the: 1) solidarity, 2) specialisation, and 3) monopoly paradigms. "Each paradigm attributes exclusion to a different case and is grounded in a different political philosophy: republicanism, liberalism, and social democracy (Silver, 1994, p. 539)."

The solidarity paradigm has its roots in French Republican thought. Via this paradigm, social exclusion will occur when the social bond between society and its citizens breaks down (Silver, 1994). Social exclusion threatens and reinforces cohesion in society. Therefore, integration is the opposite of exclusion. Silver (1994) notes that assimilating into the dominant, mainstream culture is the principal solution to social exclusion for the solidarity paradigm.

The specialisation paradigm is found mostly within the United Kingdom and the United States. Silver (1994) recognises that within this paradigm, social exclusion is seen as a consequence of both the economic division of labour and social differentiation. It is assumed that people are different, which gives rise to specialisation in the labour market and within social groups. This paradigm takes an individualist approach to understand the causes of social exclusion. Silver (1994) acknowledges that the cause of social exclusion does not lie with just the individual but also with the structures that people have created via cooperation and competition. Social exclusion is then a form of discrimination as group boundaries are created that tend to impede an individual's ability to participate in various social exchanges. Silver (1994) does recognise that within the liberal states, like the United States, individual rights should be protected. It is this protection that should obstruct various forms of social exclusion.

Finally, Silver (1994) acknowledges the monopoly paradigm. This paradigm is influential amongst the European Left. Here, social exclusion arises from the coaction between class, status, and political power, which usually serves the interest of the included (Silver, 1994). Individuals in society are excluded from full participation as institutions tend to create boundaries that keep others out against their will (Silver, 1994). These boundaries result in and extend inequality. The solution to combat social exclusion in this paradigm, Silver (1994) notes, is via notions of full citizenship, ensuring that everyone has equal membership and full participation in society.

The various causes of social exclusion, as presented by Levitas (2005) and Silver (1994) offer insight into the ways in which social exclusion could be understood in a context in which it is not often used. Silver's (1994) research has highlighted that what social exclusion means will vary in different contexts due to different traditions, etc. For guidance in discussing social exclusion in this thesis, I do not follow any of the causes or discourses of social exclusion uncovered by Levitas (2005). This is because it was developed with British, not American, public policy in mind. Silver (1994), on the other hand, provides a lens into what the causes of social exclusion could mean in an American context via the specialisation paradigm.

I use the specialisation paradigm as a guide to understanding social exclusion but with some specific caveats. Firstly, this paradigm recognises that exclusion is a consequence of individual actions and structural influences. In this thesis, I submit to a structural interpretation of the causes of poverty that treats poverty as a characteristic of society, not of individuals (Beeghley, 1988). It is a social problem related to the degree of economic and social opportunity available (Tomaskovic-Devey, 1987). This understanding of poverty is chosen in contrast with American social science, public policy, and public discourse, in which the dominant perspective conceives poverty as an individual failing (Duncan, 1984; Feldstein, 1998; Gilder, 1981), neglecting the structural component of disadvantage. The focus on individual failings neglects to view poverty as a result of structural failings (Beeghley, 1988; Brady, 2009; Duncan and Tickamyer, 1988; Lichter *et al*., 2005; Rank *et al*., 2003; Wolf, 2007).

In the second instance, Silver's (1994) discussion of the specialisation paradigm has a particular focus on labour market exclusion and the difficulties of integrating society members into it (Dean, 2016). The focus solely on the labour market to address social exclusion is incomplete (Laderchi *et al*., 2003). This is a data-driven research project in which a methodological approach is selected (section 6.2) that will derive the specific factors of social exclusion for an American context from a theoretically derived social exclusion framework (addressed in Chapter 3). Similar to Laderchi *et al.* (2003), I may find that labour force participation is not the only component of social exclusion in the United States. Subsequently, I remain open to the causes of social exclusion that are not entirely related to the labour market.

## 2.4 Disadvantage among subgroups of the American population

So far, this chapter has only considered disadvantage at an overall national level. However, it is important to consider the experience of disadvantage at the individual level. The sociology literature has paid attention to an exploration of how identity can form a basis for disadvantage (Jackson, 1998; Kabeer, 2000; Purdie-Vaughns and Eibach, 2008). In this instance, an individual can experience disadvantage in society just by virtue of who they are. Such devaluation of identity can have profound effects on individual well-being (Beals and Peplau, 2005) and on their capacity for agency (Kabeer, 2005). Subsequently, there is recognition that forms of disadvantage based on identity give rise to a horizontal model of variation as the disadvantages in question cut across economically defined vertical models and differentiate the ability of different groups within society to access resources and opportunities and participate fully in society (Stewart, 2005; Langer and Brown, 2007).

Research question two enquires about the extent to which sociodemographic characteristics are associated with multidimensional disadvantage in the United States (section 1.1). Therefore, the following discussion explores the relationship between disadvantage and group membership on the basis of age, race, gender, and citizenship status (Brown, 1995; Starrels *et al.*, 1994). In addition to these characteristics, I explore intersectionality to understand how gender intersects with race to offer a unique experience of disadvantage beyond the experience based on gender or race alone. This section highlights that subgroups of the adult population can experience disadvantage differently because of their individual characteristics and group membership. This further supports the argument that an updated conceptualisation of disadvantage is needed in the United States.

### 2.4.1 Intersectional disadvantage

Typically, poverty studies within the United States have focused attention on the effects of poverty across social categories that can be generally defined by either race, gender, or age (Norris *et al.*, 2010). This focus on singular categories does make us aware that women and minorities are often much more at risk of poverty, as will be highlighted in section 2.4.2. Examining poverty via these analytical categories offers descriptive value but often leaves the discussion of those in poverty as an issue mostly experienced by 'women & minorities.' Framing discussions in this way is a part of the 'ampersand problem,' recognised by Bowleg (2012). The issue with the ampersand and the studies that use it is that it neglects a missing piece of the intersection between multiple social identities and consequently, how that intersection might influence an individual's experience of multidimensional disadvantage. Research in public health (Bowleg, 2012), education (Núñez, 2014), and consumer research (Saatcioglu and Corus, 2014) have highlighted that social categories, such as race and gender, do coexist and depend on each other for meaning. The concept of intersectionality allows for an exploration of how identities coexist and impact individual disadvantage.

Intersectionality is a Black Feminist theory coined and articulated by Crenshaw (1991) that seeks to address the lives of individuals who suffer multiple disadvantages but are not explicitly attended to (and often excluded) by the social movements committed to overcoming disadvantage. Intersectionality has emerged as a major research paradigm that examines the multiple overlapping layers of marginalisation that affect an individual's life (Crenshaw, 1991, 1989). This subsequently has had a profound effect on studies of social group representation, because it seeks to explore how certain identities can only be understood within the narratives of other identities (Yuval-Davis, 2006). There is a key recognition in intersectional frameworks that members of disadvantaged groups are not homogenous (Bowleg, 2008). For instance, Crenshaw (1991) acknowledged that many Black women in the United States face degrading cultural representations in regards to being Black (their race), but also some face gender-related domestic and/or sexual abuse. Intersectionality provides, then, a number of strategies to explore the differences and the similarities within and across various disadvantaged and marginalised groups that experience an intersection of characteristics. Subsequently, intersectionality aids us in grasping how social identities, such as race and gender, intersect at the micro-level of individual experience to reflect interconnecting systems of oppression, disadvantage, and privilege at a more macro social-structural level (Bowleg, 2012; Crenshaw, 1991; Davis, 2008).

Adopting an intersectional approach to analysing any social phenomena requires the researcher to be reflexive of their own position in the context of analysis (Hankivsky and Christoffersen, 2008; Hunting *et al.*, 2015). I am a Black American woman. I, like Crenshaw (1989, 1991), am interested in the experience of Black women in the United States, because there remains a gap in the literature that highlights their experiences and challenges. Consequently, there is an intentional focus on the intersection between race and gender in this study that is in line with Crenshaw's (1991) original articulation of intersectionality. This study continues her original focus and makes a special case to interpret the story of Black women in the United States.

An important feature of intersectionality to note is its focus on class. Crenshaw's (1989, 1991) articulation of intersectionality showed how our understandings of social location can be improved by acknowledging how the systems of race, class, and gender overlap. Indeed, being a part of a lower social class in addition to being a minority woman may present a different social location than a minority woman with a more privileged class status. Considering health as an example, social determinates of ill health for a Black (race) mother (gender) living in income poverty with diabetes may function differently than a White father in poverty with the same health condition. Our understandings of race and class are gendered and can operate to produce varied outcomes (Crenshaw, 1991). Indeed, Caiola *et al*. (2014) acknowledge that race and gender can be "classed" just as gender and class can be "raced."

Class in intersectional frameworks have been articulated as living in poverty (for instance, Caiola *et al.*, 2014). Because this thesis is an exploration of a reconsidered look at poverty in the United States, I focus on an intersectional analysis that incorporates race and gender for reasons stated earlier in this section.

Intersectionality is helpful in understanding how individual identities relate to disadvantage. At the individual level, I am employing descriptive intersectionality, in which intersectional heterogeneity in disadvantage outcomes is explored (Bauer and Scheim, 2019; Hancock, 2007). Bauer and Scheim (2019) acknowledge analytical intersectionality that can be utilised to understand how processes, such as social exclusion, may cause a variation in outcomes. In the following chapter, I will address other forms of intersectionality, particularly structural and political. These can be analysed and examined in tandem with social exclusion. I will highlight how intersectionality, like social exclusion, considers the intertwined economic, social, cultural and political contexts in which individuals and external conditions interact (Saatcioglu and Corus, 2014).

### 2.4.2 Singular categorical disadvantage

In analysing intersectionality, the singular categories that make up the intersection are still considered (Dubrow, 2008; McCall, 2005). Therefore, gender disadvantage and racial disadvantage are both analysed in this thesis. In addition, I am also concerned with age disadvantage and citizenship (immigration) disadvantage. This section addresses these sociodemographic characteristics. In examining these sociodemographic characteristics, I am able to make the first attempt at exploring their relationships with social exclusion in the United States.

***Gender disadvantage***

Much thinking about income poverty is often gender-differentiated (Jackson, 1998; Pearce, 1993). It becomes difficult not to differentiate it as the literature provides evidence suggesting that women tend to have higher instances of income poverty than men in the United States, 14.8% compared to 12.2% in 2015[9] (Proctor *et al*., 2016). Additionally, the percentage of women in income poverty has increased since the Financial Crisis (Engle, 2013). Proctor *et al*. (2016) acknowledge that the gender differences in income poverty are much more pronounced for working-aged women, where the income poverty rate for this subgroup was estimated at 14.2% compared to 10.5% for men in the same category. These differences in income poverty rates have led to conversations on the

[9] In total number, this translates to over 24 million women and 19 million men considered income poor in the United States in 2015 (Proctor *et al*., 2016).

feminisation of poverty (Chant, 2003; Peterson, 1987; Starrels *et al.*, 1994) as was first discussed by Pearce (1978). By including a gender component to the analysis here, I am testing if there are different associations of multiple forms of disadvantage between men and women beyond income poverty. Because social exclusion is characterised as a process of disadvantage (de Haan, 1999), I am contributing to the literature via the exploration of how gender differentiates the social processes that lead women to poverty, which Razavi (1998) suggests captures the essence of gender poverty analysis.

***Racial disadvantage***

The United States operates as a racialized social system, where various levels of society, whether it be economic, political, or cultural, are structured by placing individuals into racial categories (Bonilla-Silva, 1997; Hunt and Megyesi, 2008). Politically and economically, Worts *et al.* (2010) recognise that race is one of the areas of social isolation that continues to play a key role in defining poverty trajectories. Indeed, Proctor *et al*. (2016) note that in 2015, the poverty rate for non-Hispanic Whites was lower than any other racial group in the United States. For minorities in the United States the poverty rate was an estimated 24.1% for Black Americans, 21.4% for Hispanics and 11.4% for Asians (Proctor *et al*., 2016). In analysing race in this thesis, I am testing if racial differences in disadvantage extend beyond what is available about economic disadvantage in the form of the official poverty measure.

***Citizenship disadvantage***

Immigration trends of the tail end of the 20th century produced dramatic changes in the ethnic composition of the American population (De Jong and Madamba, 2001; Suro *et al*., 2011). About 30 years ago, the foreign-born population of the United States was estimated at 7.9% (Waters and Eschbach, 1995). In 2015, 13.5% of the population were foreign-born (Proctor *et al*., 2016). With the increase in the foreign-born population in the United States, there have been renewed discussions about the integration of immigrants into American society (Garcia, 1981; Rodriguez, 1999; Treas and

Mazumdar, 2002). Most often this research has focused on the economic standing of immigrants (De Jong and Madamba, 2001). For instance, Raphael and Smolensky (2009) recognise that income poverty rates amongst the poor tend to be higher for immigrants than non-immigrants. This can be exacerbated or mitigated by the length of time an immigrant is in the United States. In comparison, Gee *et al*. (2016) acknowledge that citizenship does provide advantages, including access to the labour market, educational attainment and higher income. This suggests that it is worth exploring the relationship between citizenship status and other forms of advantage that extend beyond income. Therefore, I will explore the differentials between various types of citizenship status in the United States and multiple forms of disadvantage at the state level (further addressed in section 2.5).

***Age disadvantage***

Age, in poverty studies, is often explored via two different avenues. Firstly, age is divided into categories, in order to discuss how different age groups experience poverty and why they experience it in such a way. Age is then typically divided into the following groups: children (individuals aged less than 18)[10] (Brooks-Gunn and Duncan, 1997; Garmezy, 1993; Korbin, 1992; Korenman *et al.*, 1995; Smeeding and Thévenot, 2016), individuals of retirement age (individuals aged 65 and up) (Walker, 1980; Willson and Hardy, 2002), and individuals of working age (individuals between the ages of 18 and 64) (Brady *et al.*, 2013). For individuals over the age of 65, the emphasis in much of the literature suggests that poverty is a condition that is triggered by retirement, widowhood, and other adverse economic developments later in life (Dannefer, 2003), which tends to be felt more by women (Minkler and Stone, 1985). Amongst the working-age population, evidence suggests that working individuals represent an expanding portion of the income poor in the United States (Brady *et al*., 2013). The economic standing of the working poor can be characterised as unstable and insecure (Lee *et al*., 2005). Exploring age and its relationship to disadvantage is an important line of examination because

[10] Children are noted to be the largest and fastest growing poverty group in the United States (Brady, 2004; Korbin, 1992; McCarty, 2016). I recognise the importance of child poverty, but due to limitations in the data for this group (further discussed in Chapter 5), I do not move forward with the discussion specifically on child poverty.

each age group sheds light on the multifaceted nature of disadvantage and how experience may be varied amongst the population. Secondly, age, in poverty studies, has been discussed over the life course (Ferraro and Shippee, 2009; Rank and Hirschl, 2001). In this case, there is an exploration of the dynamism of income poverty and analysis of if, and how, people move in and out of poverty during their lives (Bane and Ellwood, 1983; Cellini *et al.*, 2008). As will be addressed in Chapter 4, this thesis uses one year of data (2015) from the American Community Survey and is thus unable to explore long term multidimensional disadvantage. Therefore, it is beyond the scope of this thesis to examine life-course patterns of social exclusion in the United States. The examination of age in this thesis is a contribution to the literature because I am able to explore age categories for multiple forms of social exclusion, which has not, to the best of my knowledge been analysed in the United States. From the above discussion, we have evidence of categorical differences between different age groups in income poverty, but not for other disadvantages. This thesis fills this gap in the literature.

## 2.5 Contextual heterogeneity in the individual experience of disadvantage

Section 2.4 acknowledged the variation within and across subgroups of the United States population. However, patterns of disadvantage are shaped by individual as well as contextual factors. The creators of the SPM (Citro and Michael, 1995) have even recognised this as they geographically adjust needs (section 2.2). In this section, I explore how individual disadvantage can be shaped by the context within which social and economic processes take place. In this thesis, place, the American states, is the context of interest. There is interest in place as a context, because space may be an important dimension in structuring multidimensional disadvantage (Buck, 2001). The United States presents a compelling case because sub-nationally, there are 51 relatively independent governments (the 50 states and the District of Columbia[11]). States define spatial and other boundaries of inclusion

[11] The District of Columbia (D.C.) is not technically one of the 50 states. I, however, lump it with the other states, thereby recognising it as the 51$^{st}$ state. This is done for the following reasons. In the first instance, the District

or exclusion. On the basis of those boundaries, states confer citizenship and rights to individuals within those boundaries (Leyshon and Thrift, 1995). In addition, there is already substantial cross-state variation in income poverty (Laird *et al.*, 2018; Triest, 1997), suggesting that it is worth exploring if there is variation among multiple disadvantages, particularly as there is much variation between the states in policies and institutions that may influence individual disadvantage (Brady *et al.*, 2013). In answering research question three acknowledged in section 1.1, I am able to explore the variation in multidimensional disadvantage across the states.

Two types of literature support the examination of state-level variation in disadvantage: comparative institutions and American states as polities (federal devolution of powers) (Brady *et al.*, 2013; Reeves, 1990). The comparative institutions' literature demonstrates that institutions and power relations organize the distribution of resources. For example, Kleiner and Ham (2002) note that industrial relations institutions affect income distribution and voice in the political system, which are noted to be important elements in the analysis of social exclusion (Labonté *et al.*, 2011). In addition, the literature suggests that economic institutions are also crucial for resource distribution (Acemoglu *et al.*, 2004). As a result, some groups or individuals in society are able to gain more benefits than others. In analysing social exclusion, this points to the idea that exclusionary processes can occur at both the micro and macro levels.

The literature on comparative institutions is only relevant here because of the role of the state in the United States. Under the federal system of the United States, there is significant devolution of powers to states, particularly for social and economic policies (OECD, 2000). States often implement and control the policies and institutions relevant to the analysis of social exclusion. For example, on specific issues such as education, health and hospitals, crime, and transport, the state governments

---

of Columbia (also referred to as Washington, D.C.) does have many similar relationships with its citizens as do the other states. Secondly, in the data used for this analysis (addressed in Chapter 4), Washington, D.C. is treated as a state. Finally, in much research on poverty and in analysing cross state differences, Washington, D.C. is treated as the other states, particularly in quantitative research where the data are available (for instance, Brady *et al.*, 2013).

are responsible for its function and implementation (Bwire and Rodríguez-Pose, 2003). The state can prevent the occurrence of risk factors associated with poverty and minimise the penalties associated with poverty. For instance, the state can aim to reduce risk prevalence of being in poverty by promoting education or employment (Laird *et al.*, 2018).

The literature on the federal devolution of powers reinforces this notion that American states are polities. Reeves (1990) notes that American states are the primary domestic governors, who make major policy decisions, are charged to protect public health and safety via the provision of various services, such as Medicaid, among other things. In addition, the states establish policy priorities and often resolve decisive social issues (Reeves, 1990). As a result, states have become more relevant settings for the struggles and settlements over the distribution of resources (Brady *et al.*, 2013). As section 2.3 highlights, these are important components of social exclusion. Social exclusion can occur at various levels, which makes the state an important level of analysis in measuring and understanding exclusion.

Consequently, where an individual resides can have a substantial influence on their experience of disadvantage. In this study, I consider place as an important factor in understanding individual disadvantage. While examining the variation in an outcome across states has been conducted before (Brady *et al.*, 2013; Osypuk *et al.*, 2006; Sampson and Lauritsen, 1997), none explored social exclusion in a multilevel context. The examination of contextual variation in the individual experience of multiple forms of disadvantage is then a contribution of this thesis (analysed and presented in Chapter 9).

## 2.6 Conclusion

This chapter addressed the key limitations of the current poverty measure in the United States and argued that there is a need for a reconsidered understanding of disadvantage. One of the primary limitations of the measure is that it solely considers money income in determining poverty status. In filling a conceptual gap in American poverty literature, I advocate for the conceptualisation of disadvantage as social exclusion. Social exclusion addresses many of the measure's limitations, notably, it considers multiple social and economic manifestations of disadvantage.

In recognising the various reasons illuminating why the American measure is not fully capable of capturing the multiple facets of disadvantage, I highlighted how various individual characteristics and the context in which an individual resides could shape disadvantage. The one size fits all approach to understanding disadvantage currently used in the United States is unable to handle this complexity. The following chapter will provide an analytical framework to measure social exclusion that is informed by intersectionality that can be explored in an American context.

# Chapter 3

## An intersectionality-informed analytical framework of social exclusion

### 3.1 Introduction

In the previous chapter, I acknowledged that the concept of social exclusion could address many of the limitations of the traditional conceptualisation and measurement of poverty in the United States. For instance, whereas the American measure focuses solely on low levels of income, social exclusion is a multidimensional concept. There is a recognition of different forms of disadvantage. People can be disadvantaged in more than one dimension at the same time. Utilising this concept allows me to move away from a consistent focus on income in American poverty studies (for instance, Brady *et al.*, 2013; Brown and Hirschl, 1995; Plasman and Rycx, 2001). It is used in this thesis in order to quantify multidimensional disadvantage in the United States, as acknowledged in Chapter 1 to be the main aim of this thesis.

The purpose of this chapter is to present an analytical framework for defining and quantifying social exclusion that is informed by the literature discussed in section 2.3. I divide this chapter into several sections. Firstly, section 3.2 discusses this research project's focus on social exclusion as an outcome. Section 3.3 introduces the framework used to measure social exclusion at the individual level, the Bristol Social Exclusion Matrix (B-SEM). Next, in section 3.4, I extend current understandings of social exclusion with incorporations of intersectionality. As an analytical tool, intersectionality is much more than a descriptive analysis of intersectional positions at the individual level (Bauer and Scheim, 2019; Bowleg, 2008; Cho *et al*., 2013; Crenshaw, 1991; Hancock, 2007). It is also a framework that allows us to examine the various structures of exclusion that can lead to varied outcomes at the individual level (Cho *et al*., 2013). Like social exclusion, intersectionality considers the entwined economic, social, cultural, and political contexts in which individuals and structural conditions interact (Saatcioglu and Corus, 2014). Therefore, I enhance this framework and apply an intersectional component to its framing. Section 3.5 concludes the chapter.

## 3.2 Social exclusion as an outcome and a process

In presenting this analytical framework, I recognise that social exclusion has been classified as both an outcome and a process (Galabuzi, 2016). As an outcome, its multidimensionality is key. As a process, social exclusion is concerned with institutional rules and relationships that distribute resources and assign value in society (Kabeer, 2000). Additionally, it implies that one type of disadvantage can lead to other disadvantages (Lin and Harris, 2008). This thesis is particularly concerned with social exclusion as an outcome (the consequences of these processes) in its application to an American context. This is for two reasons. Firstly, I recognise that using standard income-based frameworks to represent disadvantage is inadequate to understand the many dimensions of hardship. An important contribution of this thesis is the application of a concept (social exclusion) that can capture multiple forms of disadvantage in a context in which the concept has not been used. This thesis seeks to determine what the various disadvantages are in the United States. This work subsequently opens the door to future research in understanding the processes of American social exclusion. Secondly, process tends to imply something occurring over time. Indeed, there is recognition that individuals or groups who are socially excluded are often historically and systematically denied access to various rights, opportunities, and resources (Bebbington *et al*., 2008; Kabeer, 2000). This gives the impression that the quantification of social exclusion should be conducted utilising longitudinal data in order to explore the dynamic nature of disadvantage (Barnes, 2005). As will be discussed in Chapter 4, I am conducting this empirical analysis of social exclusion using a single year of data. As a result, I am unable to explore social exclusion over a long period of time. A focus on social exclusion as both process and outcome is beyond the scope of this research.

Therefore, the analytical framing of social exclusion in this thesis is focused on the concept as an outcome. As will be addressed in the next section, there have been studies focused on social exclusion as an outcome, so I am confident that this focus is adequate.

## 3.3 Analysing the outcome(s) of social exclusion: The Bristol Social Exclusion Matrix (B-SEM)

The starting point for this framework is to define what is meant by social exclusion. As was acknowledged in the previous chapter, there is no agreed-upon definition of social exclusion. There are, however, recognised components of social exclusion that a definition and subsequent framework should recognise and include. The literature agrees that social exclusion is multidimensional, relational, dynamic, relative, and encompasses agency (Barnes, 2005; Burchardt *et al.*, 2002; Room, 1995). Keeping in mind the key characteristics of social exclusion, I have selected to utilise (and adapt where noted) the Bristol Social Exclusion Matrix (B-SEM), developed by Levitas *et al.* (2007). The B-SEM defines social exclusion as

> "...a complex [outcome and] [12] process operating across several dimensions or domains. It involves both the lack or denial of resources, rights, goods and services and the inability to participate in the normal relationships and activities available to the majority of people in a society, whether in economic, social, cultural or political arenas. It affects both the quality of life of individuals and the equity and cohesion of society as a whole" (Levitas *et al.*, 2007, p. 86).

Table 3.1 presents the B-SEM, which depicts social exclusion as operating within three interconnected domains: resources, participation, and quality of life. This taxonomy is not definitive, but it was developed with the dynamic nature of social exclusion in mind. The domains are considered relevant to the four stages of the life course: childhood, youth, working-age adulthood, and later life.

[12] This thesis defers slightly from Levitas *et al.*'s (2007) definition, identifying social exclusion as outcome and process.

As a theory-driven framework, the B-SEM was derived without any reference to existing datasets, which suggests that it can be broadly applied to the analysis of social exclusion in various contexts.

**Table 3.1: Bristol Social Exclusion Matrix (B-SEM) for measuring social exclusion**

| Domains | Subdomains |
|---|---|
| ***Resources*** | Economic resources<br>Social resources and capital<br>Access to public/private services |
| ***Participation***[13] | Economic participation<br>Cultural capital and participation<br>Political and civic participation |
| ***Quality of Life*** | Health and well-being<br>Living environment & standard of living<br>Crime<br>Transportation |

Adapted from Levitas *et al.* (2007)

As far as I am aware, the B-SEM has not been applied to American data, thus is a contribution of the thesis.[14] It has, however, been empirically used in papers for the United Kingdom Cabinet Office that focused on exploring social exclusion among British families with children (Oroyemi *et al*., 2010) and young people (Cusworth *et al.*, 2009). Outside of this, the framework has been used by Crous and Bradshaw (2017), who adapt it to operationalise child social exclusion with data from 16 countries, not including the United States. Finally, Main and Bradshaw (2014) also assess childhood

[13] An aspect of the 'participation' domain within the B-SEM that is not included in this framework is social participation. Levitas *et al.* (2007) note the dividing line between the various forms of participation (social, cultural, political, and economic) are difficult to distinguish. For instance, activities such as engaging in work may or may not involve social interaction. Levitas *et al*. (2007) do suggest 'participation in common social activities' and 'social roles,' but what those actually mean are hard to define. Therefore, I choose not to include this subdomain.

[14] Relatedly, I am not aware of any study that empirically examines social exclusion for an American context.

disadvantage in the United Kingdom utilising the B-SEM. The B-SEM has also been used in studies within volumes 1 (Dermott and Main, 2017) and 2 (Bramley and Bailey, 2018) of Poverty and Social Exclusion in the UK. For instance, in the first volume, Patsios (2017) utilises the B-SEM to explored differences in social exclusion among older adults and pensioners in the UK. In the second volume, Bailey *et al.* (2018) utilise two different methodological approaches (an index and factors) to explore social exclusion utilising UK data and to explore the differences between these approaches.

Digging more in-depth at how this framework has been applied in those studies, I find that my focus on social exclusion as an outcome is sufficient. The studies by Crous and Bradshaw (2017) and Main and Bradshaw (2014) both look at social exclusion at one point in time. The others (Cusworth *et al*., 2009; Oroyemi *et al*., 2010) explore social exclusion over the life course. Because the B-SEM has been applied using longitudinal and cross-sectional data, I am confident that it can be applied to the cross-sectional data used for this analysis (discussed in Chapters 4 and 5) to explore the various forms of disadvantage at the individual level.

Like Crous and Bradshaw (2017), I accept the subdomains (and the indicators to represent those subdomains addressed in Chapter 5) of the B-SEM prima facie.[15] As acknowledged above, it was derived after an extensive review of the social exclusion literature and without any reference to a particular country context. A key objective of this thesis is to apply the concept and framework to a context in which it has not been used before. Therefore, instead of relisting the various components of the B-SEM domains, I address the strengths and limitations of utilising the B-SEM.

### 3.3.1 Strengths and limitations of the B-SEM in this context

In accepting the domains and indicators of the B-SEM (addressed in Chapter 5), I recognise the framework's strengths and limitations, particularly as I use it for this thesis. Other studies operationalising social exclusion that have not applied the B-SEM tend to use dimensions labelled

[15] There are minor changes made to names of the domains. These changes are referenced within this chapter and Chapter 5.

‘economic,’ ‘social,’ and ‘institutional’ (e.g., Burchardt *et al.*, 1999; Pirani, 2011). These domains are adequate for the analysis of social exclusion, as it does refer to economic, political, and social spheres of life. However, these dimensions separate the various components of exclusion without recognition of how the various domains can intersect. A strength of the B-SEM is that it will allow for a methodological approach (discussed in Chapter 6) that simultaneously considers the meaning and the consequences of multiple categories of disadvantage without separating the economic and social indicators that can influence and comprise each domain. Via the use of this framework, I am attempting to understand how these dimensions depend on one another for meaning and how their joint association influences the outcome of exclusion. To the best of my knowledge, Levitas *et al*. (2007) did not explicitly consider intersectionality as a particular strength of the B-SEM. As acknowledged above, it was derived after a review of the literature on social exclusion, which does not explicitly refer to intersectionality.

There are three related limitations that I address, which I see as further contributions of this thesis. Firstly, a limitation of the B-SEM relates to its application to an American context. The discussion of the B-SEM is concentrated on British and other European data. This is an expected limitation because social exclusion is an under-researched concept in the United States. The American literature is primarily limited to authors acknowledging that poverty should be conceptualised as such or acknowledging that more research on the concept is needed (Besharov and Couch, 2009; Blank, 2008; Brady, 2003b; Glennerster, 2002; Johnson and Mason, 2012; Silver and Miller, 2003). Therefore, this thesis is essentially testing if the framework could be relevant to a different country context and to see if the groupings identified in Table 3.1 hold despite the different context.

Secondly, I may find that I am unable to include each domain of the B-SEM in my analysis. I am using a data set (discussed in Chapter 4) that was not designed with the analysis of social exclusion in mind. Subsequently, it is possible that my analysis, though empirically driven may miss key components of what social exclusion entails. However, this thesis is a starting point to understanding American social exclusion and in some ways, exploratory in nature, because I am arguing that disadvantage in the United States is not limited to a lack of income as it is officially recognised. Should the analysis in Chapter 7 uncover multiple dimensions of disadvantage using this framework, even with limited indicators, then a key objective of this thesis is reached.

Finally, a limitation of this framework is that there is not an explicit focus on intersectionality. There is not an explicit recognition that the various forms of exclusion can intersect with one another and lead to varied experiences of exclusion at the individual level. In response, I have imparted intersectionality on to this framework in two ways: 1) by utilising a methodological approach (discussed in Chapter 6) in which I am testing if the discrete grouping of the domains in the B-SEM hold true or intersect for an American context and 2) by explicitly including an intersectionality informed enhancement to this framework (discussed in section 3.4).

Taken together, I find the B-SEM to be a useful framework for the purposes of this research project. It provides a sound basis on which to build understandings of the concept of social exclusion in the United States.

## 3.4 Intersectionality-informed understanding of social exclusion

The B-SEM framework reflects a substantial potential shift in thinking about multidimensional disadvantage in the United States. This can be critical in expanding conceptions of poverty beyond individual and household level features toward acknowledging the roles played by structures, institutions, and social relations (Hunting et al., 2015). As acknowledged, intersectionality allows us to examine the various structures of exclusion that can lead to varied outcomes at the individual level (Cho *et al.*, 2013). In fact, Crenshaw (1994) identified forms of intersectionality - structural, political, and representational - that make moving beyond descriptive intersectionality a worthy line of investigation.

Structural intersectionality reflects the ways an individual's social needs can marginalise them because of the convergence of multiple identities (Crenshaw, 1994; Shields, 2008). Put another way, structural intersectionality refers to the burdens faced by people who experience multiple forms of disadvantage, compounded by structural discriminatory practices (Crenshaw, 1994; Lee and Brotman, 2013). Crenshaw (1991) offered the experience of women in battered women's shelters as an example. Many of the women who seek help in these cases often are unemployed or underemployed and income poor. Crenshaw (1994) asserts that shelters have to address the violence these women have faced, along with the multiple layers of domination that converge in their lives that keep them in the less than ideal situations that brought them to the shelter. Women of colour are often more burdened by income poverty, lower levels of in-demand job skills, and childcare responsibilities (Crenshaw, 1991). When exacerbated by racially motivated discriminatory practices in the employment and housing markets, it is possible to see how economic considerations can lead to class structuring that place women of colour in some of the most disadvantaged of situations (Crenshaw, 1991; Elliott and Joyce, 2004; Massey and Denton, 1993; Peterson and Krivo, 1999).

As Crenshaw (1991) uses structural intersectionality to indicate how the interaction between various disadvantages are directly relevant to individual experience, political intersectionality indicates how the intersections of various disadvantages are relevant in the political arena (Egumenovska, 2012). Crenshaw (1991) acknowledges that political intersectionality highlights how women of colour are positioned within at least two subordinated groups that tend to pursue two conflicting agendas in the political arena. Consider that racism, as experienced by men of colour, has determined the issues of antiracist strategies; similarly, feminist, anti-sexist strategies have been determined based on the experience of White women (Crenshaw, 1994; Hall, 2015). Crenshaw (1991) asserts that the problem is that the anti-racist and anti-sexist agendas singularly do not acknowledge the extra burden of racism or patriarchy faced by women of colour. Women of colour do not experience racism the same as men of colour, nor do they experience sexism as White women may.

> "The failure of feminism to interrogate race means that the resistance strategies of feminism will often replicate and reinforce the subordination of people of colour, and the failure of antiracism to interrogate patriarchy means that antiracism will frequently reproduce the subordination of women. These mutual elisions present a particularly difficult political dilemma for women of colour" (Crenshaw, 1991, p.1252).

Finally, representational intersectionality addresses how the intersection of characteristics influence and enforce oppressive stereotypes that can preclude individuals from full participation in society (Crenshaw, 1991). In various forms of media, White women are often represented in a variety of roles, often positive, but these varied roles are not extended to women of colour. For instance, Black women have been represented in the media in more narrow, often stereotypical roles as sassy, wild, criminals, and hyper-sexualised (Crenshaw, 2014; hooks, 1992). These depictions of women of colour tend to influence support for public resources and welfare (discussed later in this section).

While considering the various forms of intersectionality as recognised by Crenshaw (1991), I now present some intersectional enhancements to the analytical framing of social exclusion that explicitly highlights the role of social locations, structural characteristics, and multilevel contexts in shaping social exclusion. To the best of my knowledge, intersectionality has only informed the measurement of social exclusion in the work of Hunting *et al*. (2015). In their framework, they assess how an intersectionality enhanced framework of social exclusion influences health outcomes in Canada. As far as I am aware, an intersectionality informed analytical framework of social exclusion used in order to assess multidimensional disadvantage has not been applied for the United States. This is, therefore, a contribution of this thesis.

The intersectionality enhanced framing of social exclusion in this thesis is predicated on three generally recognised understandings of social exclusion acknowledged by and adapted from Hunting *et al*. (2015, p. 109):

- Individuals can experience social exclusion and inclusion simultaneously.
- An individual's experience of social exclusion can be varied based on identifying characteristics.
- Social exclusion is constituted and shaped by structures of power at multiple levels.

### 3.4.1 Individuals can experience social exclusion and inclusion simultaneously.

In many studies, income poverty is constructed as a binary outcome of poor vs. not poor (e.g. Brady *et al*., 2013, 2009; Brady and Burroway, 2012). Social exclusion, in contrast, is a multidimensional concept with various indicators and manifestations. Therefore, it should not be conceived as a "binary and polarised formulation of inclusion and exclusion" (Jackson, 1999, p. 132). A separation of this concept into a binary of inclusion and exclusion ignores the fact that many individuals may be excluded in one domain and included in another. Conceiving social exclusion as a binary outcome would neglect to consider that various systems of power - like discrimination and racism - that can shape exclusion in one domain from some people across populations whilst shaping inclusion for others at the same time (Hunting *et al*., 2015). By rejecting a binary conceptualisation of social exclusion, I am able to employ a method that considers simultaneously and independently the various domains of the B-SEM showcased in Table 3.1.

This highlights the relational nature of social exclusion. A relational understanding of both intersectionality and social exclusion highlights that various forms of disadvantage can be related (Kerner, 2012). There are various systems in place that converge to shape experiences of privilege and penalty between and among groups. Baca Zinn and Thornton Dill (1996) recognise that there is comprehensive agreement that intersections can create opportunity and oppressions. An individual who is on the right side of advantage can utilise their access to status and opportunities that might not be available to one who is disadvantaged due to a particular intersection (Shields, 2008).

### 3.4.2 An individual's experience of social exclusion can be varied based on identifying characteristics.

I acknowledged in the previous chapter (section 2.4) that amongst subgroups of the American population the experience of disadvantage can be varied, which necessitates an updated conceptualisation of disadvantage that does not rely on one indicator. A benefit of the B-SEM is that it is considered relevant to the various stages of the life course. Even then, the individual experience of disadvantage can vary across the life course due to other identifying characteristics beyond age, including race, gender, and citizenship status, particularly in a diverse population as found in the United States. An intersectionality informed model of social exclusion recognises that individual experiences of social exclusion vary across multiple social locations and identities (Hunting *et al.*, 2015). Therefore, this thesis rejects relying solely on singular social categories of identity to understand the experience of social exclusion in the United States. Instead, descriptive intersectionality, in which I analyse the intersection between race and gender[16] at the individual level, is also employed.

There are two noteworthy outcomes associated with conceptualising individuals who experience social exclusion as mutually constituted by intersecting categories of identity. Firstly, I make no assumption that putting emphasis on categories like 'the excluded' or 'the disadvantaged' is sufficient for understanding the experience of social exclusion at the individual level. This is particularly important because social exclusion has various manifestations and these terms are limiting because what they mean can vary across time and space (Hunting *et al.*, 2015). Secondly, this conceptualisation explicitly recognises and rejects how categories and labels have served as tools of power and oppression (Hunting *et al.*, 2015). For example, the policies designed to help the income-

[16] The choice to analyse this particular intersection of characteristics was acknowledged in Chapter 2 (section 2.4.1).

poor in the United States have been driven by the intersection between race and class (Quadagno, 1994). In addition, Gilens (1999), implicitly recognised representational intersectionality, acknowledging that animosity toward racial minorities explains much of the American reluctance to support welfare and help the income poor. This means that not only are racial minorities facing disadvantage economically but also politically, which has consequences in other aspects of their life. Recognising the limitations of singular categorical approaches to understanding poverty is essential, because without it, the complexities of people's identities, social and historical contexts, like the influence of racial and lower class animosity on policy, are overlooked (Hunting *et al.*, 2015; McCall, 2005).

### 3.4.3 Social exclusion is constituted and shaped by structures of power at multiple levels.

Social exclusion is a concept that encompasses agency (Atkinson, 1998; Saunders, 2013). It examines who is being excluded and who is doing the excluding. Cultural capital, a subdomain of the 'participation' domain of the B-SEM, is an important component of social exclusion as it provides a focus on the excluders (Brown, 1995; Lamont and Lareau, 1988; Levitas, 2004). Cultural capital is intrinsic to the creation of a barrier between the rich (who in some instances, are the excluders to cultural participation) and everyone else in society (Levitas, 2004). For Bourdieu (2008[1986]), cultural capital is about the ways in which classes are structured and the income rich preserve power and privilege (Lamont and Lareau, 1988; Levitas, 2004). Therefore, in order to adequately address this key feature of social exclusion, an intersectionality informed framework has to acknowledge the diverse social locations and identities and how they might be situated in power relations (Hunting *et al.*, 2015). This means that the complexity of the interplay between age, geographic location, gender, and race in shaping the individual experience of disadvantage is recognised. Therefore, I examine these intersections and locate where individuals with these characteristics are positioned in various dimensions of disadvantage in relation to other individuals (Chapter 8).

The B-SEM recognises the importance of interpersonal networks in the analysis of combating social exclusion (Levitas *et al*., 2007). An intersectionality informed analysis recognises that all relationships in society are shaped and informed by a socio-political setting that either promotes inclusion or undermines it (Hunting *et al*., 2015). It explicitly recognises that there are mutually constituting structures that shape socio-political elements, further shaping boundaries of belonging (Hunting *et al*., 2015). This encompasses structural intersectionality.

The B-SEM, while an adequate framework to analyse social exclusion, does not explicitly acknowledge the intersections between various structural characteristics and their subsequent influence at the individual level. The intersectionality enhancement to the social exclusion framework explicitly recognises the intersectional structures that determine and shape social exclusion and create intersectional outcomes at the individual level (Cho *et al*., 2013; Crenshaw, 1991; Dill, 1983; Hunting *et al*., 2015). I offer two examples to help illustrate this point. Firstly, consider that via an exploration of crime, a component of the quality of life domain in the B-SEM, [17] it is possible to explore the social conditions that make it likely to commit crimes without contributing to an individual blame argument of crime, similar to Blau and Blau (1982). This focus on conditions is consistent with the social exclusion literature (Foster, 2000; Gould *et al*., 2002; Pain, 2000; Seddon, 2006). Gould *et al.* (2002), for instance, implicitly recognised that the intersection between labour market disadvantage (part of the 'participation' domain) and income poverty (part of the 'quality of life' domain) have a positive relationship with crimes committed by low skilled men. Secondly, consider racial disadvantage, which Levitas *et al.* (2007) acknowledge as a risk and driver of exclusion, and segregation. While I explore race at the individual level and not as a component of social exclusion, it still helps in understanding how race can be used as an advantage. In a racialized social system like the United States (Bonilla-Silva, 1997), White individuals are much more likely to use their privileged position in society to reside in more advantaged neighbourhoods (Krivo *et al*., 2009). Minorities, on

[17] I have titled this subdomain 'crime,' instead of 'crime, harm, and criminalisation' as termed by Levitas *et al.*(2007).

the other hand, are least afforded this privilege, in part due to discriminatory housing market practices (Peterson and Krivo, 1999). Subsequently, they are more likely to live in more disadvantaged communities, thereby bearing the brunt of crime (Krivo *et al.*, 2009; Shihadeh and Flynn, 1996). The relational nature of social exclusion and intersectionality is again highlighted here. Social exclusion is a condition where individuals are disadvantaged because social relations prevent them from improving their condition (Moncrieffe, 2004). Disadvantage, then, is a consequence of historically developed economic and political relations (Mosse, 2007).

Consequently, an intersectionality focused analysis of social exclusion has to pay attention to the interconnected nature at multiple levels of analysis (Engberg-Pedersen *et al.*, 2010; Hunting *et al.*, 2015). In recognising that social exclusion can occur due to processes occurring at various levels, it is concerned with the vertical relations between the macro and micro levels and the horizontal relations that occur within and between households, cities, and states (Engberg-Pedersen *et al.*, 2010). Similarly, intersectionality is concerned with understanding the effects of various levels in society (Hankivsky *et al.*, 2012).

To apply this enhancement to the analytical framework of social exclusion, I focus on two levels of analysis: 1) the individual and 2) the American state.[18] Firstly, the individual is the unit of analysis in this study. The analysis of poverty typically happens at the household level. Indeed, the United States officially determines income poverty status based on household statistics. This focus is insufficient because using the household as a level of analysis fails to capture intra-household differences in resource allocation (Vijaya *et al.*, 2014). These differences in resource allocation tend to differ along gender lines. This is recognised by Gornick and Jäntti (2010), who note that exploring gender differentials in poverty leads to methodological issues because gender is an individual characteristic and income poverty is, for the most part, a household concept. It becomes essential in this analysis to measure disadvantage at an individual level to circumvent these issues. Following

[18] I further discuss the reasoning for examining the individual and the state in Chapter 5 (section 5.4).

Millar's (2003) recommendation, this framework adopts a gender-sensitive analysis by looking inside the household and examine social exclusion at the individual rather than household level. Secondly, I explore variation in disadvantage across states for reasons specified in Chapter 2 (section 2.5). Subsequently, I am able to explore social exclusion vertically and horizontally.

The intersectionality-informed enhancements to measuring social exclusion offer several benefits. Firstly, I explicitly recognise that structural characteristics intersect to create disadvantages at the individual level. Secondly, I recognise that the intersection of characteristics at the individual level can result in different experiences of disadvantages. Thirdly, I acknowledge that a multilevel analysis that incorporates context is crucial to fully understand intersectionality and the experience of social exclusion. An analysis that incorporates a framework that examines intersectionality at the individual and structural levels is acknowledged to exemplify intersectional research (Bauer and Scheim, 2019; Bowleg, 2008).

## 3.5 Conclusion

The purpose of this chapter was to present an operational framework for which to empirically examine multidimensional disadvantage in the United States via the lens of social exclusion. The B-SEM, a theoretically derived framework developed by Levitas *et al. (*2007) is used to accomplish the objective set forth for this thesis. The contribution of this thesis is to apply this framework into a context in which disadvantage equates to low income and explore what multidimensional disadvantage in the United States looks like. Maintaining this focus in tandem with the characteristics of social exclusion addressed in this chapter helps to ensure that this thesis's operationalisation of social exclusion does not fall back on deficit theories that focus on the behaviour of individuals or groups of individuals (Leney, 2013), which have prominence in much American discourse.

# PART 2: DATA AND METHODS

The aim of this thesis is to use the concept of social exclusion to quantify multidimensional disadvantage in the United States. By exploring the insights social exclusion can provide in understanding the factors influencing individual social and economic disadvantage, this thesis answers the following research questions:

1. What is/are the factor(s) of multidimensional disadvantage in the United States? (Chapter 7)
2. To what extent are sociodemographic characteristics associated with multidimensional disadvantage in the United States? (Chapter 8)
3. Is there variation in multidimensional disadvantage, on average, across the United States and the District of Columbia? (Chapter 9)
    - Does that variation, if any, still persist after controlling for individual characteristics?
    - Does the relationship between individual sociodemographic characteristics and multidimensional disadvantage vary significantly across U.S. states and the District of Columbia?

In Part 1, I have explored the literature from which this thesis draws and to which it contributes. I discussed: the limitations of the official measure in the United States, the history and discourses of social exclusion, disadvantage amongst subgroups of the American population, and contextual heterogeneity in the individual experience of disadvantage. In addition, I have discussed the analytical framework that will guide the measurement of social exclusion. This framework provides the grounding for the analysis throughout this thesis, particularly for the selection of American indicators (Chapter 5) and the selection of a methodology (Chapter 6) in order to uncover dimensions of disadvantage in the United States (Chapter 7).

In Part 2 (chapters 4-6), I acknowledge the data and the methods used to answer the research questions listed above. In chapter 4, I will discuss the data used to answer the research questions, the 2015 Public Use Microdata Sample files of the American Community Survey. I will also address the strengths and limitations of using that data to achieve the research aims. In chapter 5, I operationalise social exclusion. I recognise the variables from the data used to measure social exclusion. Finally, in chapter 6, I discuss the three methods (factor analysis, ordinary least squares multivariate regression analysis, and multilevel modelling) I use to answer the research questions listed above. Additionally, I address the implications of using these methods with the data discussed in Chapter 4.

# Chapter 4

# Selecting an appropriate American data source to measure an un-American conceptualisation of disadvantage

## 4.1 Introduction

The focus in the United States on utilising income as the sole indicator of disadvantage is reflected in the limited availability of data that can fully capture the multidimensional nature of disadvantage. To reach the research aims of this thesis, it becomes necessary to look beyond traditional data sources. This thesis utilises an American government agency produced data set, the 2015 American Community Survey (ACS) Public Use Microdata Sample (PUMS) files. The ACS PUMS files are classified as 'big data,' which incorporate additional strengths and limitations associated with utilising secondary data. In this chapter, I discuss the features of the ACS PUMS data and the implications for use in this thesis. I begin by introducing the idea of big data (section 4.2). Secondly, I present the data, describing its origins and development (section 4.3). I then discuss using the data for secondary analysis (section 4.4). Next, I consider the strengths and weaknesses of the data (section 4.5). Finally, I acknowledge the ethical considerations of using this data as it relates to the research aims (section 4.6). The chapter is then concluded.

## 4.2 Big Data

The term 'big data' is a buzzword in the social sciences, with no set definition. Manovich (2012), however, notes that the term 'big data' can be applied to data sets whose size surpasses the ability of commonly used software tools to manage and process the data within a tolerable elapsed amount of time. With more than 2.3 million sample members in the 2015 ACS PUMS data set, the data used within this thesis can be classified as big data. Based on my experience with the data, the ability of software packages to manage and process the data resonates with Manovich's (2012) definition. However, Floridi (2012) notes that this definition is vague, because it should not be entirely focused on how many data we can process. Instead, Laney (2001) suggests three dimensions of data management - velocity, volume, and variety (the three V's) - that have, subsequently, been used to discuss and describe big data (for instance, Chen *et al.*, 2012; Kwon *et al*., 2014; Russom, 2011). Big data that captures 'the Vs' has been defined as ...

> "...a term that describes large volumes of high velocity, complex and variable data that require advanced techniques and technologies to enable the capture, storage, distribution, management and analysis of that information" (TechAmerica Foundation's Federal Big Data Commission, 2012, p. 10).

It is important to address and describe what is meant by the 'three Vs' and its application here. First, velocity is the rate at which data can be generated and the speed at which it could and should be analysed (Gandomi and Haider, 2015; Russom, 2011). As will be discussed later in this chapter, the creation of the ACS data allows the Census Bureau to provide information on characteristics of the American population yearly as opposed to every 10 years as was once done (section 4.3). Volume can refer to the magnitude of the data (Gandomi and Haider, 2015; Russom, 2011). When addressing the volume, there is recognition that definitions of big data volume are relative can vary by different factors, such as the type of data (Gandomi and Haider, 2015). The ACS data are collected for a sample of addresses taken for the entire American population. Over 3 million addresses are sampled each year. In Chapter 7, I will discuss how the size of the data influenced the analysis undertaken for this thesis utilising the statistical software package, Stata. Variety, Sagiroglu and Sinanc (2013) suggest, is what makes big data really big. It can come in variety of sources: structured, unstructured and semi-structured. As is discussed in section 4.3.1, the ACS PUMS are presented to the analyst relatively structured, such that the information is provided in spreadsheets and is organised so that the data can be analysed by machines (Gandomi and Haider, 2015; Russom, 2011).

In addition to volume, velocity, and variety, Gandomi and Haider (2015) recognise 'value' as an additional attribute of big data. They acknowledged that big data have been characterised as having a low value density, implying that the data in its original form has a small value when compared to its size (Gandomi and Haider, 2015). A way to combat this, they suggest, is to analyse large volumes of such data (Gandomi and Haider, 2015). I recognise the value of the ACS data beyond analysing large volumes of it. For this analysis, I use just one year of the data available and still find value for this analysis. In this chapter and the next, I highlight why and how the ACS PUMS is the best available set of American data to understand and develop dimensions of disadvantage in order to reach the primary research aim of this thesis (highlighted further in section 4.5).

## 4.3 The American Community Survey (ACS)

Prior to 2010, data users interested in gathering information on the American population for various social, economic, and housing characteristics had to utilise the long form of the decennial Census produced by the United States Census Bureau.[19] Despite its important uses, including the allocation of federal funding, the long form of the Census has had substantial limitations. Firstly, as the Census data are collected once every ten years, the information obtained tends to be outdated. For instance, MacDonald (2006) notes that decennial data were usually up to ten years out of date. Out of date information from administrative sources (such as the federal government, in this case) meant that subnational governments at the state and county levels were not effectively able to chart and predict social and economic trends (MacDonald, 2006; Van Auken *et al.*, 2006), particularly during times of rapid demographic change. Secondly, the accuracy of the Census was not definite. Often times, users

[19] The short form of the decennial Census, on the other hand, is used to count the population for a variety of purposes, including political redistricting (Van Auken *et al.*, 2006).

of Census data would only have access to population estimates based on the sample of households who actually completed the long form of the decennial Census. The response rates for the decennial Census were lower than 65 percent (MacDonald, 2006).

Recognising these limitations, the Census Bureau began developing and testing a new means of data collection via a 'rolling sample' or 'continuous measurement' survey during the 1990s (Williams, 2013). The concept of a monthly rolling survey was proposed and developed by Leslie Kish in a series of papers (Kish, 1998, 1990). In contrast to the decennial Census that collected data at one point in time, the new survey would collect data continuously from independent monthly samples of the population and aggregate the data to produce estimates. In addition, the Census Bureau wanted the survey to have many of the same features of the decennial Census, such as mandatory responses to some questions, mailed survey questionnaires, and follow-ups via telephone and in-person visits if necessary (Williams, 2013).

The reimagined survey became the American Community Survey (ACS).[20] The Census Bureau (2014) recognises three main objectives for the ACS. First, the ACS provides federal and subnational (state and county) governments with an information base for administering and evaluating government programs (Census Bueau, 2014). The second objective is to replace the long form of the decennial Census with the ACS so that the decennial Census can be used solely for counting the population (Census Burea, 2014). The third objective is to provide researchers and other data users with yearly information on comparable statistics across states, communities, and population groups (Census Bureau, 2014). It is the third objective that facilitates the ACS' use in this study.

[20] To the best of my knowledge, there are no user groups dedicated to the ACS that, for example, offer advice in analysing the data. However, the Census Bureau does allow users to share ACS data stories. There are four stories published on the site (https://www.census.gov/programs-surveys/acs/about/acs-data-stories.html) as of 15 July 2019. There is not a question and answer feature on this website. These stories mainly focus on government bodies using the data to support government projects.

The ACS was fully implemented in 2005 (Census Bureau, 2014, 2008). Full implementation sees the ACS expand to all 3,141 counties in the 50 states and the District of Columbia. Since 2011, the ACS collects data from a sample of nearly 300,000 housing units a month, totalling about 3.54 million a year (Census Bureau, 2014). The data found within the ACS are summed over time by government analysts and other researchers in order to provide large enough samples for reliable estimates. Population and housing profiles for the first year of full implementation became available in 2006 and have been available every year for geographic areas with populations of at least 65,000. Longer accumulations of data are necessary for less populous areas. Consequently, the first three-year estimate was released in 2008, which consists of combined data single-year ACS for the years 2005, 2006, and 2007. These estimates include areas with populations equal to or greater than 20,000 (Herman, 2008; Nirel and Glickman, 2009). The first five-year estimate was released in 2010 and combined ACS years 2005-2009. This includes information on the smallest geographic areas, with populations less than 20,000, in the United States. When the first five year estimates were made available, the Census Bureau was able to replace the long form of the decennial Census, as accumulations of ACS data provide products similar to the ones obtained in the long form of the Census (Census Bureau, 2018, 2014, 2008). The phases of development for ACS subsequently provided a reputable source of social and economic information for the American population without waiting ten years for the latest information.

### 4.3.1 Sample Design

People are not the unit basis of the sample for the American Community Survey; addresses are (Herman, 2008). The Census Bureau's sampling frame for the ACS is the Master Address File (MAF). The MAF is the Census Bureau's official inventory of known city and non-city mailing addresses[21] for housing units, group quarters, and selected non-residential units in the United States (U.S. Census

[21]The Census Bureau (2014) notes that a city style address is one that has a structure number and street name format, while a non-city style address is one that has a rural route or post office (PO) box number format. City style addresses are most prevalent in the United States and in 2010, accounted for 98.2 percent of the residential addresses in the MAF (U.S. Census Bureau, 2014).

Bureau, 2014). Via the MAF and in files called MAF extracts, the Census Bureau has information on mailing and location addresses, physical characteristics and/or location descriptions of units, and any other relevant attribute of each living quarter. This implies that the sampling frame for the survey captures most of the American population.

The ACS sample is based on a two-phase stratified sample designed to identify 3.54 million housing units annually. The first phase of sampling involves selecting housing unit address samples for the 3,141 counties and equivalents in the United States, including the District of Columbia (Census Bureau, 2014). This is done in two stages, which for ease are labelled here as 1a and 1b. During phase 1a, the Census Bureau assigns the addresses attained from the MAF to five equal and representative sub-frames. It is a design requirement of the ACS that addresses in one sub-frame are eligible to be in the ACS sample each year and each sub-frame is used every fifth year (Census Bureau, 2014). Essentially, one address in the United States can receive a questionnaire no more frequently than once every five years (Herman, 2008). It can be inferred that no address in the 2015 ACS PUMS used for this analysis was included in any ACS PUMS files after 2010.

Phase 1b involves selecting a sample of addresses from the sub-frame of the current year (Census Bureau, 2014). The sub-frame is divided by county and the addresses are selected from the sub-frame in each county. The second stage allocates the sample to the twelve months of the year for data collection, the process of which leads to the initial annual ACS sample (Census Bureau, 2014). Another part of this stage is to assign the sample addresses for the year to a specific data collection month while ensuring that one address does not receive several other Census Bureau surveys in the same month.

The second phase involves sampling for computer-assisted personal interviewing (CAPI) follow up (Census Bureau, 2014). Consider that the ACS is conducted as a mixed-mode household survey, in which the internet, mail, telephone, and personal visit are used as modes of data collection. Sampled households have to respond via any of these models within a three-month window, as the survey is mandatory by law (Groves, 2012). These households are first sent an ACS questionnaire by mail. Nonresponse to the mail survey results in a telephone call to complete the survey using computer-assisted telephone interviewing. This practice of successive contact results in a very high response rate of 96% (Davern *et al.*, 2009). When household units do not complete the survey via the internet, mail or computer-assisted telephone interviewing by the third month, they, along with undeliverable addresses, are eligible for CAPI (Census Bureau, 2014). At the end of each year, all the data collected from the preceding year are processed into a single data file for each year.

The extensive sample design employed by the Census Bureau (2014) has implications for this thesis. Firstly, as addresses are the basis of the ACS sample, this would suggest that the analysis within this thesis is examining the places where people are, not necessarily the people. This implies that the characteristics of each individual at the respective housing unit and any subsequent disadvantage they face may change as their addresses change. It also implies that those without addresses, such as the homeless are less likely to be counted in the analysis. Secondly, the ACS does not count people at a certain point in time like the decennial Census. It measures the population characteristics by aggregating statistics collected in each monthly survey (Herman, 2008). There is some chance, therefore, that the aggregated statistics over the course of the year include individuals that have moved. Finally, because the survey is required by law, the ACS has high participation rates (Groves, 2012). The high rates imply that some of the difficulties faced by users of the long form of the decennial census, including biased estimates due to low response rates, are not likely to be faced here.

### 4.3.2 Weighting

As is common with most household surveys, weights are used in the ACS to ensure the characteristics of the sample are in agreement with those of the full population. The Census Bureau (2014) employs a method to compensate for the differences in sampling rates across various areas, differences between the full sample and the interviewed sample, and differences between the sample and independent estimates of basic demographic characteristics. The method used is called a ratio estimation.

The ratio estimation method enables the Census Bureau (2014) to take advantage of independent population estimates by age, gender, race, and Hispanic origin, and estimates of total housing units produced by the Population Estimates Program (PEP) of the Census Bureau (2014). The PEP produces estimates of the population for the United States and demographic characteristics are produced in each of those areas as of 1 July. Because of this method, the ACS is able to produce the same demographic characteristics for the same set of statistical, administrative, and legal bodies that was previously published in the long form of the decennial Census (Census Bureau, 2014).

The use of the ratio estimation method results in the assignment of two sets of weights: 1) a weight to each sample person record and 2) a weight to each sample housing unit record (Census Bureau, 2014). For any geographic area, a characteristic total is estimated by summing the weights assigned to the people, households, and other housing units possessing the characteristics. Whereas estimates of population characteristics are based on the person weight, estimates of characteristics associated with the family, household, and housing unit are based on the housing unit weight (Census Bureau, 2014). The advantage of ratio estimation is that it allows for an increase in the precision of the estimates and corrects for under and over coverage by individual demographic detail and by geography, which is useful for this thesis in addressing the research question that assesses the relationship between sociodemographic characteristics and area disadvantage (**Does the variation in multidimensional disadvantage across the United States persist after controlling for individual characteristics?**[22]). The estimates produced from the ACS are based on a probability sample and will vary from their actual population values due to sample and non-sampling error (Census Bureau, 2018, 2014). In addition, the estimates from the ACS will vary based on the combination of interviewed and

[22] This is the first sub-question from research question 3: Is there variation in multidimensional disadvantage, on average, across the United States? I changed the wording of the sub-question slightly to fit the context here. These sub-questions are further acknowledged in Chapter 6 (section 6.4).

non-interviewed housing units in each tabulation month. Part of the process of calculating the person weights for the ACS is to control estimates of totals by age, race, gender, and Hispanic origin to be equal to population estimates by weighting area. The Census Bureau (2014) offers two reasons for this: 1) to reduce the variability of the ACS housing unit and person estimates and 2) to reduce bias due to under-coverage of housing units and the people within them in household surveys. The bias that results from missing these housing units and people is partially corrected by using these controls. Subsequently, these controls allow the sample to be more representative of the American population, the population of interest.[23] Therefore, the generalisability of the findings in this research is enhanced and more accurate results are produced when answering research question two (**To what extent are sociodemographic characteristics associated with multidimensional disadvantage in the United States?**).

### 4.3.3 Public Use Microdata Sample (PUMS) files

The data estimates generated in the ACS are not customisable. Therefore, this analysis will utilise the ACS Public Use Microdata Sample (PUMS) files. The PUMS data products are derived directly from the ACS and contain actual responses from the full survey. The PUMS files enable researchers to create their own estimates of demographic, housing, economic, and social characteristics and cover geographic areas within the United States from individual ACS records (Census Bureau, 2014, 2008). This gives me greater flexibility in constructing the relevant measures of social exclusion later in this thesis (Chapter 7).

The PUMS dataset includes variables for practically every question in the ACS, as well as many different variables that were derived from multiple survey responses, such as poverty status[24]

[23] This is likely to not include undocumented immigrants. Section 5.3.4 discusses the variable from the data set that captures citizenship status. While it is possible to determine if a respondent was born outside of the United States, it is impossible from that information to determine if they are undocumented and living within the United States.

[24] Poverty status, for instance, is determined by taking the ratio of individual income and the poverty line threshold. This is calculated for this researcher and included in the data dictionary of the ACS PUMS files, named '*povpip*.'

(U.S. Census Bureau, 2018, 2014). Each entry in the file represents a single person, or--in the household-level dataset--a single housing unit (Census Bureau, 2016b). In the person-level PUMS file, individuals are arranged into households, which makes it possible to study individuals within the settings of their families and other household members (Census Bureau, 2016b).

The PUMS files are slightly different from the core ACS files in at least four ways. Firstly, the PUMS files include a sample of addresses from the full ACS. In 2006, for instance, 1.2 million addresses of the 2 million that were sampled for the ACS were included in the 2006 PUMS (Census Bureau, 2008). Secondly, the only geographies in a PUMS file are the state and the Public Use Microdata Area (PUMA). PUMAs are special non-overlapping areas that divide a state; each PUMA contains a population of about 100,000 or more. This means that except for the PUMAs and state codes, there are no variables included identifying the major group quarter types or other sub-state (e.g. the county and/or city) information for the imputed records (Census Bureau, 2016b). Thirdly, the PUMS files are much more flexible than the aggregate ACS data available, though they tend to be more complicated to use (Census Bureau, 2016b). Working with the PUMS data generally involves downloading large datasets onto a local computer and analysing the data using statistical software. Finally, whilst recognising that all ACS responses are kept strictly confidential, various variables in the PUMS files have been amended in order to safeguard the confidentiality of survey respondents (Census Bureau, 2016b). For instance, particularly high incomes are "top-coded," meaning that the value in the data does not exceed $10 million or go below -$20,000.[25] Additionally, uncommon birthplace or ancestry responses are grouped into broader categories and the PUMS files provide a minimal set of geographic variables.

Because the PUMS files are the data sets available directly to researchers, there are several implications that will affect this analysis. Firstly, the limited set of geographic variables limits the analysis of contextual heterogeneity in disadvantage to the state. I am not able to explore any

[25] These specific numbers are acknowledged in the 2015 ACS PUMS data dictionary.

predefined geographic boundaries below the state level with this data. This is because one year of data limits population size to 65,000. These population sizes are more likely to be in larger-scale geographic areas. For the analysis undertaken in this thesis, it does not pose negative implications, because as acknowledged in Chapter 2, between state and within-state variation in disadvantage is where the interest of this thesis lies. Secondly, though the Census Bureau (2014) used a very detailed survey methodology for the ACS to correct for under- and oversampling of different populations, there may be additional sampling error that may result from the use of the PUMS. Therefore, the Census Bureau suggests the use of replicative weights which are available for the ACS for 2005 data onward. In the ACS PUMS data, there are eighty replicate weights at the household and person levels that allow the users of the data to generate empirically derived standard error estimates. The standard error of an estimate measures the variation of a test statistic across multiple samples of a given population (Census Bureau, 2014), indicating that the actual standard error of any estimate calculated from one sample can never be known with certainty (Census Bureau, 2016b). Using replicate weights allow a single sample to simulate multiple samples, 80 in this case, which therefore generates more informed standard error estimates that mimic the theoretical basis of standard errors while retaining all information about the intricate sample design of the ACS (Census Bureau, 2016b, 2014). The resulting standard errors are then used to obtain more precise confidence intervals and significance tests. As the PUMS data are derived from the full ACS, replicate weights are applied to bring estimates for the characteristics under analysis in this paper closer to those that would be found in the full ACS sample (Census Bureau, 2016b). The application of these weights is further addressed in Chapter 8 (section 8.2).

## 4.4 Secondary analysis of the ACS PUMS

In order to reach the main aim of this thesis and to answer the research questions laid out earlier in, a single-year PUMS file is used for the year 2015. As a result, I have data for person and housing unit profiles for each state, as they have populations greater than 65,000. Any estimate released in a given year is a period estimate, which is an average of data collected in every month during the previous year. For example, for the 2015 ACS, the data were issued in 2016 and represent information gathered throughout 2015, not a particular point in 2015.

The ACS PUMS is a hierarchical or nested data set. Aarts *et al.* (2014) note that data which are characterised by a hierarchical structure are organised at more than one level. In this data structure, individuals are nested within households within states. In total, the 2015 PUMS data contains 1,363,661 housing unit records, 3,028,122 person records as nested within households, and 149,650 person records for those living in group quarters (Census Bureau, 2016b).

The analysis of the American Community Survey is secondary, as the data are being repurposed (Cheng and Phillips, 2014; Hand, 2018). There are many benefits associated with utilising secondary data. One of the most noted benefits of utilising secondary data for analysis is the low cost (Cheng and Phillips, 2014). The amount of data available for analysis utilising a secondary source would be impossible for me to collect alone. Collecting data for a large scale survey is resource-intensive and can prohibit researchers from collecting primary data. The Census Bureau employs a team of data scientists, statisticians, and data collectors dedicated to capturing a vast amount of data with ready to use survey weights and design variables for researchers, like myself, to conduct new types of analyses previously unavailable. Some data sets have a fee associated with accessing it, but the Census Bureau provides open access to the ACS PUMS data sets. This allows data researchers to use existing data to test new and exciting hypotheses using various statistical models (Cheng and Phillips, 2014).

Inherent to the nature of utilising secondary data is that the data were not collected to answer the questions under investigation. The ACS PUMS data are a set of data that Connelly *et al*. (2016) recognise as 'found' data, in which the data was not collected for the purpose of research and in which researchers have no input into the design and structure, or even content of the data. There is a compromise that has to be made utilising this data set between cost and relevance. For instance, a survey could be designed specifically to measure multidimensional disadvantage via the concept of social exclusion, though costly, whereas the free, open-access data might be only roughly suitable (Hand, 2018). As collecting primary data is out with of the scope, time, and budget of this analysis, compromise is made to repurpose data. As will be discussed in the following chapter, these compromises are reflected in variable selection. Cheng and Phillips (2014) note that it is not uncommon that some variables are unavailable for analyses when utilising secondary data.

## 4.5 Strengths and weaknesses of using the ACS PUMS

When utilising data for secondary analysis, Cheng and Phillips (2014) point out that researchers should have a clear and comprehensive understanding of the strengths and weaknesses of the data set. The purpose of this section is to discuss those in regards to the ACS PUMS data. As the discussion above has begun to highlight, the ACS PUMS data set is good for the analysis undertaken within this thesis. To the best of my knowledge, the ACS PUMS offer the sole source of the sociodemographic characteristics necessary to capture the multidimensionality of the concept of social exclusion to measure disadvantage in the United States.

The suitability of the ACS PUMS data for this analysis is highlighted when it is compared to other alternative data sets in the United States. Here, I discuss the CPS primarily, because it is often used in studies dedicated to examining income poverty (Brady *et al*., 2013; Cook and Frank, 2008; Lochner *et al*., 2001). Additionally, it is the official source of income and poverty data in the United States. The CPS is recommended by the Census Bureau (2008) for use on related subject areas. I do, however, also provide an appendix (Appendix A) that lists other competing data sets that may have been used in this analysis, along with further justifications of the suitability of the ACS for this analysis. The CPS provides data primarily on the monetary aspects of poverty, such as income and employment. Subsequently, the recommendation from the Census Bureau to use the CPS on poverty-related research reflects a narrow view of poverty that is purely based on a household's level of income.

For the purposes of this thesis, the ACS has several advantages over the CPS and other data sets. Primarily, the ACS collects information on the non-economic aspects of poverty that is useful for quantifying multidimensional disadvantage. The CPS collects information solely on income and employment. In addition, the ACS benefits from a sample size that is thirty times larger than the CPS (Davern *et al.*, 2009). As acknowledged in Appendix A, the sample size in the ACS is substantially larger than other potential data sets. This suggests that the single-year estimates of the ACS at the state level will be more accurate[26] than those from the CPS and other data sets. The CPS is designed to be state representative but its sample size does not support accurate single-year comparisons for many subgroups, such as those articulated in Chapter 2 (section 2.4) across states or time. Boudreaux *et al.* (2011) acknowledge that the relative size of the ACS' state samples suggests that the single-year estimates will be sufficient for most research tasks. Furthermore, the ACS has been chosen over the CPS in a recent study on multiple deprivation in the United States (Dhongde and Haveman, 2016), endorsing the suitability of the ACS' for this thesis. These factors, along with the high response rate of the ACS, allow for an analysis of the required subpopulations across large geographic areas within the United States.

### 4.5.1 Limitations of the ACS PUMS data

The ACS is the best available data source for the needs of this thesis, as far as I am aware. However, the most notable limitation is that the ACS PUMS data were not collected with the concept of social exclusion in mind. As such, not all indicators relevant to measuring social exclusion will be addressed.

[26] While I recognise the many benefits of having a larger sample size, I do acknowledge that there is a possibility that many of the relationships I test in later chapters are more likely to be statistically significant because of the size.

Positively, many variables are available in the ACS PUMS that can be utilised to construct suitable social exclusion indicators, subsequently applied in later analysis (addressed in Chapter 5).

An additional limitation of using the ACS PUMS data in analysing social exclusion relates to its dynamism. Social exclusion is characteristically dynamic, meaning that it is not a time constant quality, but instead can be seen to change over time, with an individual moving both in and out of a situation of social exclusion (Barnes, 2005). An individual can be socially excluded if the conditions she experiences persists or worsens over time (Bellani and D'Ambrosio, 2011). Because I am only using one year of data, I am not able to explore this characteristic of social exclusion. This analysis focuses then on the dimensions of exclusion in one year and the individual experience of disadvantage in that year (2015).

## 4.6 Ethical Considerations: Confidentiality in the American Community Survey

This project uses the 2015 ACS PUMS which contain data for a sample of all housing units and provides information on the housing and population characteristics of each of these units. The Census Bureau ensures that all identifying information is removed from the files and other disclosure avoidance techniques are used to ensure confidentiality (Census Bureau, 2014). In addition, the PUMS files do not contain names, addresses, or any information that can identify a specific housing unit or person (Census Bureau, 2016b). However, I recognise that this research area is a sensitive topic because I am concerned with expanding the focus of disadvantage beyond low levels of income. The intention of my work is to determine, by using data available to researchers, if people are experiencing forms of disadvantage not captured by the official poverty measure. I am acknowledging marginalisation and disadvantage in American society. Therefore, I take considerable care in presenting the research findings in Chapters 7-9. In Chapter 8, I will acknowledge how specific sociodemographic characteristics are associated with multiple forms of social exclusion. Therefore, I am conscious of the reporting of the lives of these potentially disadvantaged groups via the dissemination of the results of this research. I treat all data from the respondents with the highest levels of respect, regardless of their race, gender, income levels, citizenship status, age, or other identifying traits, particularly those utilised to answer the relevant research questions. In addition to the efforts made by the Census Bureau to ensure the data are confidential, I will ensure that respondent anonymity is protected. I have completed the University of Edinburgh's School of Social and Political Science Research Ethics Level 1 self-audit checklist.

## 4.7 Conclusion

The purpose of this chapter was to introduce the ACS and argue that PUMS files are an appropriate data source for the analysis under question. The ACS is the largest survey in the United States with a sample of more than 3 million addresses each year. It provides detailed demographic, social, economic, and housing data and to the best of my knowledge is the only suitable source of small-area data on social and economic characteristics for the nation (Census Bureau, 2018), necessary to measure multidimensional disadvantage. The ACS PUMS is arguably the best available data source, but it does have its limitations, including the fact that it was not particularly designed for the examination of social exclusion. Despite these, the ACS PUMS data set is able to support the research aims and meet the subsequent requirements for this study.

# Chapter 5

## Operationalising an American measure of multidimensional disadvantage

### 5.1 Introduction

In Chapter 3, I discussed the Bristol Social Exclusion Matrix (B-SEM) and recognised it as an appropriate framework to measure social exclusion for the purposes of this research project. The B-SEM identifies three interconnected dimensions that are relevant to the measurement of social exclusion at the individual level: resources, participation, and quality of life. The identification of appropriate indicators or suitable proxies for these domains depends heavily on its availability in the 2015 American Community Survey (ACS) Public Use Microdata Sample (PUMS) data. The purpose of this chapter is to identify and discuss the variables selected from the data set that can adequately contribute to the measurement and analysis of social exclusion in the United States. In section 5.2, I select appropriate indicators to be applied to the B-SEM. In section 5.3, I recognise the variables representing sociodemographic characteristics that will be used in later analysis. In section 5.4, I consider the adjustments made to the analysis based on the availability of suitable indicators in the ACS PUMS. Section 5.5 concludes the chapter.

### 5.2 Operationalising the dependent variable for multidimensional disadvantage

For reasons discussed in Chapter 3, the Bristol Social Exclusion Matrix (B-SEM) is used for guidance in operationalising social exclusion for the United States. As acknowledged in Chapter 3 (section 3.3), the B-SEM understands social exclusion as operating within three interconnected domains: resources, participation, and quality of life. Table 5.1 presents the domains, along with the subdomains of the B-SEM identified by Levitas *et al.* (2007)[27] and the chosen indicators from the 2015 ACS PUMS data. I am using the B-SEM to measure exclusion at the individual level. Subsequently, the variables identified are at the individual level (except where explicitly acknowledged). Blank indicators signify that no suitable variable in the ACS PUMS file exists. For instance, there are no indicators in the ACS PUMS that can represent the crime, political participation, and social participation subdomains. Levitas *et al*. (2007) note that citizenship status can serve as an indicator of political participation. While the ACS PUMS has this indicator variable, I chose not to include it as a component of social exclusion. It has been deliberately chosen to keep demographic variables out of the development of the dependent variable. This choice is made for two main reasons. Firstly, it is not my intention to have characteristics of the individual define social exclusion in the United States. As acknowledged in chapter two, I treat poverty as a characteristic of society and not of individuals (section, 2.3.2). Secondly, the relationships between sociodemographic characteristics and social exclusion are the main focus of research question two (**To what extent are sociodemographic characteristics associated with multidimensional disadvantage in the United States**?). Like age, race, and gender, citizenship status is a sociodemographic variable of interest, as acknowledged in section 2.4. Leaving these variables out of the construction of the dependent variable allows me to answer this question later in this thesis. Following this, I discuss the selected variables.

[27] The subdomains for each of the three domains are taken from the B-SEM as developed by Levitas *et al*. (2007). For reasons specified in Chapter 3 (section 3.3), I do not adjust the B-SEM except where noted.

**Table 5.1: Indicators for B-SEM, as available in the American Community Survey and adjusted for United States context**

| B-SEM Domains | B-SEM Subdomains | ACS Indicators for EFA |
|---|---|---|
| ***Resources*** | Economic resources | Total personal income<br>Food stamps |
| | Social resources & capital | Marital status<br>Unmarried partner |
| | Access to public/private services | -- |
| ***Participation*** | Economic participation | Employment status<br>Industry of employment |
| | Cultural capital & participation | School enrolment<br>Educational attainment<br>Internet access<br>English fluency |
| | Political & civic participation | --- |
| ***Quality of Life*** | Health & well-being | Health insurance coverage<br>Disability status |
| | Living environment & standard of living | Housing quality - overcrowded<br>Income poverty status |
| | Crime | --- |
| | Transportation | Travel time to work<br>Means of transportation to work |

Adapted from Levitas *et al*. (2007)

### 5.2.1 Subdomain 1: Economic resources

Two variables from the ACS PUMS are used to represent economic resources in this analysis: 1) total individual income and 2) food stamps. An individual's income is quite likely the most obvious indicator to be included in this subdomain. While it is recognised and argued within this thesis that disadvantage involves much more than just a lack of income, Lister (2004) warns of the danger of downplaying income when describing poverty. Subsequently, I recognise the importance of including income in this analysis, mainly due to the role it plays in the United States in determining eligibility for benefits.

When discussing income as an indicator for this domain, Levitas *et al*. (2007) break it into estimated income and components of income. They acknowledge that these differences in types of income may be important as different income data may be included across various surveys (Levitas *et al.*, 2007). The ACS PUMS does differentiate between an individual's earnings and income. The ACS defines earnings as the sum of wage and salary income, including self-employment income (Census Bureau, 2008). Income, on the other hand, includes earnings, which are often the most substantial part of overall income, but also money received from retirement or any other means (Census Bureau, 2008). Because an individual's earnings are included in the income variable, a separate indicator representing 'earnings' is not included in this subdomain. The 'total person income' variable that is included in this analysis is measured as a continuous variable. Due to the Census Bureau's efforts to anonymise recognisable information, such as extremely high and low incomes, income is rounded and bottom and top coded, acknowledged in the previous chapter.

'Food stamps' is the other indicator for this subdomain. Food stamps, now more commonly known as the Supplemental Nutrition Assistance Program (SNAP), is the most extensive food assistance program (and largest in-kind benefit provided to the income poor) in the United States (Schmeiser, 2012). SNAP is designed to reduce food-related hardship for low-income individuals. This is done by providing direct support to the household so that household can purchase food and alleviate hunger (Leung *et al*., 2012; Ratcliffe *et al*., 2011). Eligibility for food stamps benefits is determined at the household level and the provision of benefits is for the household. Subsequently, the respective food stamps variable in the data set is available at the household level. Therefore, if one household or housing unit receives food stamps, each individual in that housing unit is presumed to be a recipient of food stamps. There is a limitation with this, of course, as we cannot be 100% confident that each individual within the housing unit has equal use of the food stamps. However, due to the acknowledgement in the literature that these sorts of benefits should be included in discussions of disadvantage (Citro and Michael, 1995; Ploeg and Citro, 2008), it is included in this analysis. The importance of including this variable in the analysis is that a common critique of the American poverty measure is that it does not include noncash sources of income and taxes (Cancian and Danziger, 2009). Being a beneficiary of food stamps or SNAP is an important noncash source of income. In the data set, there is a binary variable for food stamps, indicating whether a household receives this benefit or not. As food stamp receipt is determined at the household level and each individual within is presumed to benefit from it, it is a suitable variable to utilise at the individual level.

### 5.2.2 Subdomain 2: Social resources & capital

In discussing this subdomain, Levitas *et al*. (2007) acknowledge an increasing awareness of the importance of social networks for individual well-being. They focus on social support and its availability on a day-to-day basis. Due to its availability in the ACS PUMS, I focus on aspects of social resources as it relates to the family. Levitas *et al*. (2007) do acknowledge that frequency and quality of contact with family, friends, and colleagues at work can be used as an indicator. While no information in the ACS PUMS discusses the amount of time an individual spends with others, it can be assumed that those who are married may spend more time together than individuals who are not. The variables selected to serve as indicators are thus selected to serve as proxies for social connections and networks: 1) 'marital status' and 2) 'unmarried partner'. The marital status variable is categorical with five categories. Respondents can classify themselves 'married,' 'widowed,' 'divorced,' 'separated,' or 'never married.' The unmarried partner variable is also included because it can serve as a proxy for social contacts and social support. If a person lives with a family member or unmarried partner, they are likely to have increased feelings of social connectedness and benefit from the networks that result. 'Unmarried partner' is a binary variable, where yes indicates being in an unmarried partnership and no indicates not.

### 5.2.3 Subdomain 3: Economic participation

Economic participation is an important component in the analysis of social exclusion as participation in the activities that generate access to goods, services, and resources are often believed to enhance an individual's well-being (Levitas *et al*., 2007; Waddell and Burton, 2007). One of the main ways many people participate in society economically is via employment. Subsequently, the variables included serving as indicators from the ACS PUMS for this subdomain relate to employment. Two indicators from the ACS PUMS are used: 1) employment status and 2) industry of employment.

In the ACS PUMS, employment status is represented by a categorical that are grouped into four categories of employment. An individual can identify as 'civilian employed,' 'unemployed,' 'in the armed forces,' and 'not in the labour force.' These categories offer a fairly simplistic look at employment status in understanding disadvantage. Due to this, there are a few limitations. Firstly, a limitation of this variable is that it does not allow us to examine various forms of underemployment such as part-time work, and overqualified workers with jobs mismatched to their skills (De Jong and Madamba, 2001). Additionally, migrant workers may not be included. It is possible that they are counted in this variable under the 'civilian employed' category, but because the data do not specifically ask if an individual is a migrant worker, I do not know if they are included. I am only able to specify whether a person has a job or not. Secondly, individuals who are undertaking unpaid work, such as an internship, or providing unpaid care cannot be included. It would be remiss to assume that because someone is not in the formal labour market, they are socially excluded from economic participation. This, in quite a few scenarios, would not be the case. However, this again relates to the limits placed by utilising a data set not designed for the analysis of social exclusion.

In the B-SEM framework, Levitas *et al*. (2007) acknowledge the nature of working life, such as type of occupation, as a component in this domain that can be tested. For this, 'industry of employment' is included as an indicator. In the United States, the industry in which an individual is employed is an important aspect of understanding disadvantage. The industry of employment and the type of job is one of the macro-level labour market characteristics that can put people at risk of poverty (Brady, 2009). The industry of employment variable is a categorical variable with many categories. In the 2015 ACS PUMS data documentation, the listing covered six pages. To reduce the number of variables that would have to be included, these were regrouped into overall industry titles, resulting in 19 categories. These categories include agriculture, manufacturing, mining, utilities, construction, wholesale, retail, transportation, information services (like media), financial services, professional services, education, medical, social care, entertainment, other services, administration, military, or unemployed.

### 5.2.4 Subdomain 4: Cultural capital & participation

Much of the work done on cultural capital has been conducted in the education field (DiMaggio and Mohr, 1985; Robinson and Garnier, 1985; Sullivan, 2001). Subsequently, much of the discussion about the acquisition of cultural capital centres on educational attainment. The indicators used here does as well, but also includes variables to capture access to education and basic skills, which Levitas *et al.* (2007) acknowledge has an effect on employability and social participation. I use the following indicators from the ACS PUMS data: 1) educational attainment, 2) school enrolment, 3) internet access, and 4) living in an English fluent household.

There is universal agreement about the relevance of education in understanding disadvantage (Alkire *et al.*, 2010), so I have two indicators that relate to education. First, I include the variable representing educational attainment. Educational attainment can be defined as the highest level of education that a person has successfully completed and is ordered hierarchally (Lutz *et al*., 2007). For the analysis, the categories are ordered as follows: no schooling, preschool, kindergarten, elementary school, middle school, high school no diploma, high school diploma, GED or alternative credential, some college, no degree, Associate's Degree, Bachelor's degree, Master's or Professional degree, doctorate degree.

Secondly, I include a 'school enrolment' variable. While the educational attainment variable captures what level of education an individual has already achieved, this indicator signals an individual's prospects for the future. It has been recognised that education typically increases a person's quality of life, allows for better health status, contributes to enjoying more social connections, amongst several other recognised benefits (Rolfe, 2012; Ross and Willigen, 1997; Sparkes, 1999; Umberson and Montez, 2010). Those currently enrolled in education are in the process of developing the skills and connections that typically come with that pursuit of education. The 'school enrolment' indicator is a three-category variable, specifying no, has not attended school in the last three months, yes enrolled in public school or college, and yes enrolled in a private school or college or home school. These variables are divided into three binary variables representing each category.

I use two variables to reflect access to knowledge and basic skills: 1) internet access and 2) English fluent household. Internet access, Levitas *et al*. (2007) note, involves access to knowledge. In societies where important components of society involve knowledge-intensive activities, the distribution of knowledge across society is linked to stratification (Hargittai, 2003). This indicator helps us to recognise the significance of the digital divide, the phrase which has been applied to the gap present in most countries between those with ready access to digital technologies and the subsequent knowledge gained from that access and those without such skill or access (Cullen, 2001; Hargittai, 2003). Due to the amount of attention that academics and policymakers put on understanding which sectors of society have access to the internet or are internet users, access is often defined as having a network-connected machine in the home or workplace and defined in binary terms (Hargittai, 2003). The ACS PUMS variable for internet access is conceptualised in binary terms, access or not. There are implications using this variable in this thesis. Firstly, access is not synonymous with use (DiMaggio and Hargittai, 2001). The ACS PUMS does not ask specific questions about an individual's use of the internet. The access to internet variable is defined at the household level and as a result, everyone who resides at the same address will have the same response for this variable, which results in further limitation. For instance, I am not able to look at differences in access within the household. I am not able to tell who uses the internet in the housing unit the most and who uses it the least.

The final indicator for this subdomain is 'English fluency' which serves as a proxy for the basic skills identified in Levitas *et al.* (2007). Stolzenberg and Tienda (1997) argue that English fluency is a form of human capital and noted that disadvantage, particularly for minorities is conditional on the level of human capital. Indeed, Throsby (1999) recognises that there are connections between human and cultural capital. In addition, Sullivan and Ziegert (2008) conclude that a lack of fluency in English contributes to high poverty rates for Hispanic individuals.[28] The variable in the ACS PUMS data set that represents English fluency is 'limited English speaking household,' for which there are two categories. The first category represents a household in which no person in the household, 14 and over, speaks English only or speaks English 'very well.' The second category represents a household in which at least one person in the household meets that criteria. Interestingly, Dhongde and Haveman (2016) utilise this particular variable as a measure of social connectedness, which highlights that there are many interconnections between this subdomain and the subdomain for social resources. Throsby (1999) acknowledges that when relationships between individuals in society are invoked, cultural capital becomes intertwined with social capital, the concept of which is utilised to

[28] Sullivan and Ziegert (2008) focus on Hispanic immigration and poverty status. It has, however, been recognised that immigrants from other countries in which English is not the native language are likely to have issues that affect quality of life and influence functioning in settings that provide essential resources, including education (Hernandez and Charney, 1998).

identify indicators for the 'social resources' subdomain. This reiterates the interconnectedness of these subdomains and further lends weight to an argument that disadvantage is not purely economic based.

### 5.2.5 Subdomain 5: Health and well-being

Two variables available in the ACS PUMS serve as indicators for the health and well-being subdomain of the 'quality of life' domain: 1) disability status and 2) health insurance coverage. A person's disability status, be it physical or mental, is a well-recognised potential barrier to integration into society, particularly via the labour market (Barnes and Mercer, 2005; Sloane and Jones, 2012) and is often due to social stigma (Ahmedani, 2011). One could argue that if an individual has a disability or disabilities, they potentially have a relatively lower quality of life compared to individuals without a disability. Albrecht and Devlieger (1999), however, find that there is a disability paradox in which people who have disability often feel they have a high quality of life. This highlights the importance of subjective perceptions and understandings of quality of life (Kim *et al.*, 2017). Due to the nature of the data set, only objective variables can be used. Consequently, I recognise the limitation of disability as an indicator for this domain when considering that there are subjective understandings of the concept. Despite this, there remains universal acknowledgement of its role in understanding quality of life, particularly in regards to integration into society (Ahmedani, 2011; Barnes and Mercer, 2005). Therefore, it is included here.

It is recognised that analysing disability can be challenging because there remains a need to develop a robust and standardized definition of it. I am bound to the definition of disability in the ACS PUMS, which is identified as severe difficulty in one of four areas of basic functioning: vision, hearing, cognition, and ambulation (Weathers II *et al*., 2005). Four binary variables, representing the absence or presence of the respective disability exist in the data set. However, in lieu of having four separate binary variables for each functioning included in the ACS PUMS, one binary variable is created. An individual having serious difficulty in one of the four areas discussed above is categorised as disabled.

There are known limitations of the ACS' definition of disability. It does not include other aspects of disability that may influence an individual's ability to participate in the normal activities of American society. For instance, the ACS does not allow for the identification of the prevalence of specific health conditions, such as cancer, paralysis, and HIV/AIDS (Weathers II *et al.*, 2005), conditions which may exclude people from society. Also, group quarters populations are not included in this analysis, which may include people with disabilities who live in group homes. The intersection between living in a group quarter and having a disability may be an important risk factor for social exclusion that cannot be included in this analysis because of this specific limitation of the ACS data.

The second indicator for the health and well-being subdomain is 'health insurance coverage.' Health insurance is an important indicator for health and well-being in the United States particularly at the time of writing this thesis. Throughout the Obama administration (2009-2017) and since health care has been at the centre of policy debates. Interestingly, in a list of possible indicators for the respective domain, Levitas *et al*. (2007) do not include health insurance. This is likely due to the UK having a single-payer, publically administered healthcare system although many employers do offer private health insurance (Haven *et al.*, 2013). In the United States, there is not a single nationwide system of health insurance. Insurance is purchased in the private markets or provided by the government to certain groups, such as for individuals with low incomes (Ridic *et al.*, 2012). An individual is able to purchase private insurance from numerous for-profit commercial insurance companies or from non-profit insurers. Ridic *et al*. (2012) acknowledge that most of the health insurance coverage in the United States is employment-related, largely due to the cost savings associated with group plans that can be purchased via an employer. This suggests that health insurance is an important indicator for this domain because it not only serves as a proxy for economic security, as acknowledged by Dhongde and Haveman (2016), but it signals an individual's ability to protect themselves from a preventable disability had they not been able to obtain health insurance (Drake *et al.*, 2009).

Individuals who lack health insurance often receive less medical care than those with health insurance, including screening and treatment (Adler and Newman, 2002). In addition, those without health insurance may also receive lower-quality care. Because health insurance, generally, is so important to improving access to health thereby improving health outcomes (Hoffman and Paradise, 2008), this thesis does not differentiate between the types of health insurance an individual has, whether it be from an employer or the government. Consequently, the thesis represents health insurance coverage via a binary variable.

### 5.2.6 Subdomain 6: Living environment & standard of living

In this subdomain, Levitas *et al.* (2007) focus on housing quality and neighbourhood quality and satisfaction. The ACS does not collect information on neighbourhood quality or satisfaction. Consequently, no indicator related to an individual's neighbourhood is included in this analysis. This presents an issue for this analysis because stratification by place is quite prevalent in the United States (Massey, 1996). Additionally, social characteristics can vary widely and systematically across communities (Sampson, 2003), particularly along dimensions of socioeconomic status, including poverty (Stafford and Marmot, 2003) and racial and ethnic composition (Massey and Fong, 1990). The indicators for this subdomain will, therefore, focus on housing quality.[29] The ACS PUMS indicator that will represent housing quality is overcrowded housing.[30] As the ACS PUMS does not directly ask respondents about overcrowded housing, this variable is constructed utilising other available variables in the data set.

Guidance on constructing overcrowding is taken from a 2007 report by the United States Department of Housing and Urban Development (2007). In the report, the authors discuss three

[29] The ACS PUMS also has the following relevant housing quality variables: complete kitchen and complete plumbing facilities. However, within the 2015 data, less than one percent of the addresses in the sample had no access to complete kitchen and plumbing facilities. Therefore, I do not include it in the analysis.
[30] Housing costs as percentage of household income was also considered to represent housing quality. It is not used here as it is endogenous with income. Individual income is included as indicator for the 'economic resources' subdomain.

common measures of overcrowding in the household: 1) the number of persons per room, 2) the number of persons per bedroom, and 3) square footage per person (U.S. Department of Housing and Urban Development, 2007). Following Dhongde and Haveman (2016) and Schill *et al.* (1998), overcrowding is measured via the number of persons per room option. Using the number of persons per bedroom would neglect to account for people who may sleep in non-traditional areas of the home, such as basements and living rooms. Overcrowding is subsequently defined, in this thesis, as more than one occupant per room. To achieve the indicator for overcrowding, the number of people in each housing unit is divided by the number of rooms in the unit. A binary variable is then created to represent overcrowded living spaces if the individual lives in a housing unit with more than one person per room.

In this subdomain, a standard of living component is also included to explicitly recognise its relationship to quality of life (Schalock, 2004). An income poverty threshold is the indicator typically used to measure standard of living (Slesnick, 1991). The suitability of its use, it is recognised, remains up for debate. Wang *et al*. (2016), however, note that an income poverty line well captures the monetary aspects of poverty and that under normal circumstances, an increase in income tends to somewhat improve well-being in all domains. Subsequently, I use an income poverty measure as an indicator here. In the ACS PUMS, official income poverty is measured at 100% of the federal poverty line. The limitations of the official poverty measure are noted in Chapter 2 (section 2.2). Therefore, it is an indicator I will not use for the standard of living subdomain. A more relative measure of income poverty is adopted, similar to the 50/60% of median income threshold often used in European and comparative poverty analysis (e.g., Osberg and Sharpe, 2014). Income poverty for this analysis is measured at 250% of the federal poverty line. Individuals and families who are between 100 and 250% of the income poverty level often face economic insecurity which is typically not captured utilising just the official American poverty measure (Kearney *et al.*, 2013). Therefore, via a binary variable, any individual whose income is 250% of the poverty line and below is categorised, in this thesis, as in relative income poverty.

### 5.2.7 Subdomain 7: Transportation[31]

Two variables from the ACS PUMS are selected as indicators for this subdomain: 1) 'means of transportation to work' and 2) 'travel time to work'. The means of transportation by which an individual travels to work is an important indicator for quality of life because work is where many people receive their income and interact with social networks (Matthews and Besemer, 2014). In addition, travelling via modes other than walking requires money and people who can afford faster modes such as the car or public transportation can reach wider opportunities that contribute to quality of life (Titheridge *et al*., 2014). The 'means of transportation to work' variable in the ACS PUMS is a categorical variable with twelve categories, including car, truck, or van; bus or trolley bus; streetcar or trolley car; subway or elevated; railroad; ferryboat; taxicab; motorcycle; bicycle; walked; worked at home; and other method. Titheridge *et al*. (2014) note that one of the resources required to travel include time. For this reason, travel time to work is included in this analysis as an indicator for this subdomain. 'Travel time to work' is a continuous variable, ranging from 0 minutes for individuals who do not work and for those who work at home to 200 minutes. The variable is top coded at 200 minutes. There are two implications for this. Firstly, due to confidentiality issues, the creators of ACS keep variables like this top-coded as extremely long travel times may be readily identifiable. This leads to the second implication. Individuals who do travel longer to work do not have their travel time listed. Subsequently, their experience travelling to work and the actual time it takes them to travel to work is not fully captured in the data.

[31] Levitas *et al.* (2007) includes transport as an indicator for the 'access to public and private services' domain. I deviate slightly from that and put transportation as subdomain of quality of life for two reasons. Firstly, transportation is recognised to be important to 'quality of life' (Titheridge *et al.*, 2014). Secondly, the other indicators noted by Levitas *et al*. (2007) relate to services available within the home, such as access to public utilities. I acknowledged in the 'living environment' subdomain that a small percent in the sample are without those services (less than 1%) and I do not include it in the analysis.

## 5.3 Identifying the independent variables

Research question two examines the extent to which sociodemographic characteristics of the individual is associated with multidimensional disadvantage. In Chapter 2, I acknowledged some of the sociodemographic characteristics that confound the effects of poverty and disadvantage in the United States. The following subsections discuss the measurement of these variables within the ACS PUMS: age, gender, race, and citizenship status.

### 5.3.1 Age

The ACS PUMS is a cross-sectional data set, which means that it is not possible to track an individual's experience of disadvantage over the life course. Instead, I am able to explore age differences in social exclusion using a single year of data for 2015. The PUMS data provide the ages for each individual in their respective housing unit, producing a variable with a continuous level of measurement. I, subsequently, group and analyse ages into different categories. I follow Dhongde and Haveman's (2016) grouping of age into six categories. Firstly, individuals aged 18-24 are grouped together, representing young adults, an age range that has garnered much coverage in discussions of various forms of disadvantage (Park *et al.*, 2006). Next, the typical working-age group is broken down into 4 groups: 25-34, 35-44, 45-54, and 55-64. This method is utilised because there is enough heterogeneity within the working-age population, warranting a separation of the categories in lieu of examining this age group homogenously. Finally, individuals aged 65 and up are grouped together to represent retirement age. [32] Grouping the age variable will allow an analysis that explores the average differences in multidimensional disadvantage that each age group may face.

---

[32] It is recognised that the literature on labour market typically leave out the category for those over the age of 65. However, as I highlighted in the previous chapter, the labour market is not the only dimension of disadvantage and the research may highlight that those in this age group may experience some disadvantage and/or exclusion that extends beyond the labour market.

### 5.3.2 Race

There are 7 categories for race: White alone,[33] Black alone, American Indian, Asian alone, Native Hawaiian/Pacific Islander, some other race alone,[34] and Mixed Race.[35] I recognise that ethnicity is an important discussion on minority disadvantage. However, the ACS PUMS does not separate race from ethnicity. For instance, there are black Hispanics in the data set and no particularly easy way to separate them without either double counting or dropping individual observations. This would bias later results. Therefore, a Hispanic origin variable is not included in the analysis, as particularly relevant for answering research question two (**To what extent are sociodemographic variables associated with multidimensional disadvantage in the United States?**). It is hoped to capture some of the relationships between ethnicity and disadvantage with the addition of the foreign-born variable discussed in section 5.3.4.

### 5.3.3 Gender[36]

In the ACS, a variable exists for gender. The gender variable is coded as either male or female. It does not specify any other category for gender identification. Subsequently, my analysis is limited to a gender discussion that includes male and female. Any disadvantages that may be associated with non-binary or transgender individuals are not included in this analysis.

### 5.3.4 Citizenship status

The variable for citizenship status in the ACS PUMS is divided into five categories: 1) born in the U.S., 2) born in U.S. territories, including Puerto Rico, Guam, the U.S. Virgin Islands, and the Northern Marianas, 3) born abroad of American parent(s), 4) U.S. citizen by naturalisation, and 5) not a U.S. citizen. I keep these categories just as they are in the ACS PUMS data. I do not create a binary variable

---

[33] In the ACS PUMS, alone specifies that the person is not of mixed race.
[34] The ACS PUMS data dictionary does not specify what these other races may be.
[35] Mixed race means that the individual identifies as having more than once race.
[36] The ACS PUMS data dictionary defines this as gender. I will follow suit, forgoing any conversation on the differences between gender and sex.

for U.S. citizen or not, because I am concerned with the differences in various forms of disadvantage between each of these categories, just as I am for age or race.

### 5.3.5 Intersectionality

In Chapter 2, I noted the importance of intersectionality in understanding disadvantage in the United States. In tandem with the recognition of its importance, there is a recognised complexity associated with measuring intersectionality (Bowleg, 2008; Dubrow, 2008; McCall, 2005). Intersectionality has a long history of application, but as Cho *et al*. (2013) acknowledge, as an analytic sensibility, its optimal application in quantitative research is quite unclear. A challenge in addressing intersectionality in quantitative research is that the ACS PUMS was not developed with intersectionality in mind. Subsequently, I am limited to the demographic categories included within the PUMS. In order to manage the complexity of measuring intersectionality quantitatively, I adopt the intercategorical approach recognised by McCall (2005). This approach requires the adoption of existing analytical categories in order to document relationships of disadvantage among social groups. McCall (2005) recognise two other approaches: intracategorical and anticategorical. Intracategorical complexity focuses on social groups at neglected points of intersections, those whose identities cross traditional constructions, whereas anticategorical complexity deconstructs analytical categories (McCall, 2005). In this thesis, I focus here on the intersection between race and gender, because this is the intersection of identity used by Crenshaw (1991) in articulating intersectionality. Additionally, these are existing analytic categories in the data and across poverty studies, facilitating an easier interpretation of results.  Consequently, the categories for race and gender will remain as they have been discussed in previous sections. Construction of the intersection between these categories is discussed in Chapter 6 (section 6.3).

## 5.4 Recounting the sample members based on available indicators

In section 4.4, I acknowledged the number of sample members available in the 2015 ACS PUMS (3,028,122 person records nested within 1,363,661 housing unit records,) After considering the relevant indicators and availability in the PUMS for the analysis of social exclusion, it was found that not all sample members available could be included in this analysis. This is for one primary reason: lack of information for relevant indicators. Because there is a lack of relevant indicators for individuals under the age of 18,[37] they are not included. In addition, individuals living in group quarters, including nursing homes, prisons, college dormitories, juvenile institutions, and emergency and transitional shelters for those experiencing homelessness are not included. There are two implications for this research project. Firstly, my analysis only explores multidimensional disadvantage for the adult, non-institutionalised population in the United States. Secondly, particularly as it relates to the group quarters populations, my analysis does not include the groups who are likely to face exclusion in various domains like the elderly in nursing homes and the incarcerated. It could be implied that my analysis might not include the dimensions of social exclusion that are particularly relevant to these groups. This is a recognised consequence of utilising secondary data. Popay *et al.* (2008) note that the use of secondary data can themselves be exclusionary because the people most severely affected by exclusionary processes – such as the homeless and the institutionalised – are often the least likely to be included. Given these considerations, I am left with 2,348,374 sample members for analysis, which can still be classified as big data and is generally large enough to allow for generalisability.

[37] Though the B-SEM can be applied to various stages across the life course, including childhood, it would still have to be adapted to fit the experience of children, as was done specifically for children by Crous and Bradshaw (2017).For this thesis, this additional analysis and adaptation of the B-SEM is not feasible given time and resource constraints.

## 5.5 Conclusion

The purpose of this chapter was to provide a list of the variables used from the ACS PUMS to measure social exclusion as recommended by the Bristol Social Exclusion Matrix. The indicators chosen from the PUMS files are those noted in the literature to capture the nature of social exclusion (Gordon, 2000; Labonté *et al.*, 2011; Levitas *et al.*, 2007). In the latter part of this chapter, I also identified the variables that represent the sociodemographic variables of interest that will aid in answering the research questions of this thesis. Finally, I considered the implications of not having available indicators for some subgroups of the American population and the adjustments made as a result.

The next chapter is the final chapter in the 'Data and Methods' section of this thesis. I will discuss the methodological approaches undertaken to answer the research questions listed earlier in this thesis.

# Chapter 6

# Methodology: Using multivariate big data to analyse multidimensional disadvantage in the United States

## 6.1 Introduction

The research questions outlined in Chapter 1, in addition to substantive concerns, encompass three methodological objectives. The purpose of this chapter is to discuss the statistical methods selected in order to achieve these objectives, particularly as is necessary and available for the analysis of 'big data.' The first objective is to construct a measure of social exclusion using the observed variables identified in Chapter 5 as indicators for the various domains of the B-SEM. Section 6.2 argues that the most appropriate method to construct social exclusion is Exploratory Factor Analysis (EFA). The second objective is to test and assess the relationship between individual sociodemographic characteristics and multidimensional disadvantage. Section 6.3 discusses how Ordinary Least Squares (OLS) multiple linear regression models can be applied to reach this objective. Finally, the third objective is to assess whether there is any state-level variation in multidimensional disadvantage across the United States. Section 6.4 argues the benefit of applying multilevel modelling techniques to the American Community Survey (ACS) Public Use Microdata Sample (PUMS) data to answer the relevant research question that explores state-level variation in disadvantage across the United States. Section 6.5 concludes the chapter.

## 6.2 Constructing measures of multidimensional disadvantage using factor analysis

Social exclusion is a latent construct, meaning that it is only possible to measure it indirectly. Its construction requires a methodological approach that allows for many or few variables to be retained as indicators of the various forms of disadvantage. For this purpose, factor analysis is utilised. Factor analysis is a broad term representing statistical techniques that allow for defining and estimating an unobserved structure underlying the variations of observed variables and their interrelationships (Fabrigar and Wegener, 2011; Hair Jr. *et al*., 2010; Matsunaga, 2010). It becomes a useful tool for investigating complex concepts that are not particularly easy to measure by minimising a large number of variables into a few interpretable factors (Costello and Osborne, 2005). It will be used here to highlight various forms of disadvantage.

### 6.2.1 The case against developing an index

Given the importance of measuring multidimensional disadvantage in this analysis and developing robust dependent measures for use in subsequent analysis, I considered developing an index. Developing an index has been used as an instrument to quantify multidimensional poverty (Myles and Picot, 2000). When indices are constructed, several factors are weighted together, as seen with the Human Development Index and the more recent Multidimensional Poverty Index (Alkire *et al.*, 2010). In essence, the same could be done with social exclusion. The preparation of a composite index for social exclusion would call for a choice of appropriate weights for the economic, social, and cultural components of the domains of social exclusion recognised in Chapters 3 and 5. The usefulness of these types of indices is questionable as it is not always sensible to aggregate various indicators, particularly when they do not move in the same direction. Importantly, building an index does not directly measure an underlying construct in question. Diener and Suh (1997) acknowledge that when indicators are combined, the general index has the advantage of simplicity and breadth at the cost of more detailed information. Using an index means that there is the potential to overlook important differences on specific indicators. This technique might not allow for an analysis of the intersections between the dimensions of social exclusion. Furthermore, certain statistics, such as shared variance, is not used in the construction of the index, which Rippin (2011) acknowledges is one of the main methodological weaknesses of the Multidimensional Poverty Index. The index that remains is a single, not multiple, construct which provides a simple summary measure that can be difficult to interpret substantially. As I am hypothesising that disadvantage is a multidimensional phenomenon, the use of a single index would essentially turn disadvantage into a unidimensional issue. Instead, factor analysis is favoured.

### 6.2.2 Exploratory factor analysis

Within factor analysis, two methods exist to derive factors from many variables: exploratory and confirmatory. Both methods are used to examine the underlying factor structure of data, but they play different roles in terms of the purpose of the given research. The exploratory factor analysis (EFA) is used for theory building and the confirmatory factor analysis (CFA) is for theory-testing (Costello and Osborne, 2005; Matsunaga, 2010). For this thesis, I use exploratory factor analysis (EFA). There are several reasons for this choice. Firstly, there is a recognition that the dimensions and subsequent indicators in this analysis are quite interconnected. It may be found that indicators theorised to be under one domain may actually group elsewhere. As such, the EFA will allow the data to dictate the dimensions under which each selected indicator fall. CFA does not let the data identify or discover the underlying dimensions (Ferguson and Cox, 1993; Hoyle, 2000; Matsunaga, 2010; Mueller, 1996; Yong and Pearce, 2013). Secondly, because social exclusion is a concept not widely studied in the United States, the use of exploratory factor analysis is better suited as the context change may challenge the existing framing of the dimensions of disadvantage analysed via social exclusion. Finally, as there are no predetermined criteria regarding the distribution of variance onto the factors, CFA would not be an appropriate factor analysis method, because with that method, the researcher knows the structure of the latent construct (Doyle *et al*, 2011; Hanna *et al.*, 2011;

Weizmann-Henelius *et al*., 2010). I am not sure which structure social exclusion would take with the ACS PUMS data. EFA is best used when there are no expectations about the underlying structure (Fabrigar and Wegener, 2011). Matsunaga (2010) notes that EFA is best used when the researcher is unsure of the principal mechanisms of the underlying concept and unsure of how the variables used for analysis operates in relation to each other.

### 6.2.3 Factor analysis' suitability to the research

Factor analysis is suitable for my research questions and the 'big data' used in this thesis for several interrelated reasons. Firstly, the variables selected to measure social exclusion, outlined in Chapter 5, are based on a theoretically derived framework, the Bristol Social Exclusion Matrix, developed by Levitas *et al*. (2007). Subsequently, it is anticipated that the variables used to measure the concept will have a common underlying construct. Fabrigar and Wegener (2011) support this reason, advocating to use factor analysis when the research is concerned with determining the number of factors a set of measured variables are assessing. Secondly, the concept of social exclusion is 'complex' and by its very nature 'multidimensional,' which are the type of data that factor analysis is suited to analyse. Next, a key aim of factor analysis is data reduction (Floyd and Widaman, 1995; Hutcheson, 1999). There are 16 variables from the ACS PUMS data theoretically identified to make up social exclusion (see Table 5.1). They need to be reduced into a smaller number of explanatory constructs so that future analysis may be possible. The statistical basis for factor analysis is that with a group of variables, parsimony can be achieved by explaining the maximum amount of common variance using the smallest number of explanatory constructs that represent dimensions in the data (Floyd and Widaman, 1995; Hair Jr. *et al*., 2010; Yong and Pearce, 2013). The resulting factors are then interpreted according to statistical and substantive criteria. Lastly, factor analysis is a key statistical technique in this analysis because its results will substantially inform each research question. The results of the factor analysis will allow me to answer research question one (**What is/are the factor(s) of multidimensional disadvantage in the United States?**). The factors derived are subsequently the dimension(s) of multidimensional disadvantage in the United States. For research question two (**To what extent are sociodemographic characteristics associated with multidimensional disadvantage in the United States?**), I examine the relationships between individual characteristics and the derived factors using Ordinary Least Squares regression analysis (discussed in section 6.3). Finally, for research question three (**Is there variation in multidimensional disadvantage, on average, across the United States?**), contextual heterogeneity in multidimensional disadvantage (the derived factors) are examined using multilevel modelling (discussed in section 6.4). Subsequently, it serves as the basis for the other statistical techniques utilised in this analysis.

In addition to ensuring that factor analysis is suitable for the research questions, the data itself has to be suitable for this technique (Fabrigar and Wegener, 2011). EFA is based on correlation matrices. Positive correlations will exist between measured variables because these variables are influenced by one or more of the same unobservable constructs (Fabrigar and Wegener, 2011). Pearson's correlation is usually used to determine if there are adequate relationships between variables. To use that method, the variables have to be measured at a continuous or interval level, though ordered categorical variables can also be used (van der Eijk and Rose, 2015). The variables identified in Chapter 5 have varying levels of measurement, as is the case with most social science data. Only two of the variables selected, 'travel time to work' and 'total income,' have a continuous level of measurement. The other variables are categorical, mostly unordered. There have been developments in statistical methodology that allow for factor analysis to be applied to variables with varying levels of measurement. In these instances where the variables used to construct the latent variable take on ordinal, nominal, and continuous levels of measurement, a polychoric correlation is advised (Holgado–Tello *et al.*, 2010). van der Eijk and Rose (2015) acknowledge that the polychoric correlation has been shown to approach the real underlying structure for these types of variables better than Pearson's correlation. Therefore, the analysis favours a polychoric correlation over Pearson's correlation.

## 6.2.4 Stages of conducting factor analysis

Once satisfied that there is sufficient correlation among variables, there are several critical decisions to be made: 1) selecting the estimation method, 2) deciding the number of factors to retain, 3) selecting an appropriate rotation method, and 4) interpreting the results of the output (Costello and Osborne, 2005; de Winter and Dodou, 2012; Floyd and Widaman, 1995). Many options are available for each step and their suitability depends on the nature of the research. This section acknowledges the decisions made for this particular analysis.

***Estimation method***

Firstly, the researcher has to settle on an estimation method. The two most common techniques are maximum likelihood and principal axis factoring (PAF). The selection of either method depends on the purpose of the research. Yong and Pearce (2013) recognise that maximum likelihood tries to analyse the greatest possibility of sampling the observed correlation matrix. PAF, on the other hand, is a least-squares estimation that seeks the least number of factors that can account for the correlation amongst the variables (de Winter and Dodou, 2012). The factors are basically extracted successively until there is a large enough of variance accounted for in the correlation matrix (Yong and Pearce, 2013). I choose to use PAF over maximum likelihood for two specific reasons. Firstly, maximum likelihood is recognised to be more useful for CFA rather than the EFA used in this analysis (Yong and Pearce, 2013). Because CFA is a method to test theories, maximum likelihood allows the researcher to test the statistical significance of the correlation among factors and the factor loadings (discussed later in this section) (Fabrigar and Wegener, 2011). Secondly, maximum likelihood requires multivariate normality (Costello and Osborne, 2005; de Winter and Dodou, 2012). When this normality assumption is violated, maximum likelihood tends to produce Heywood cases, which are invalid outcomes that occur when the factor loadings exceed 1.0 (Costello and Osborne, 2005). Three tests were conducted to check for multivariate normality: 1) Shapiro Wilk (Royston, 1992), 2) the

Shapiro-Francia (Royston, 1983), and 3) and Doornik-Hansen (Doornik and Hansen, 2008).[38] Each test was statistically significant, p<0.001, indicating that the data violate the null hypothesis of normality. Therefore, I follow Costello and Osborne's (2005) recommendation to utilise PAF.

***Factor retention***

Another essential decision to make when conducting factor analysis is determining how many factors to retain. The objective of this step is to identify the number of factors or latent dimensions needed to take into account the shared variance among the variables (Reise *et al.*, 2000). When the EFA is running, it will extract as many factors as there are variables. Each factor will load in decreasing order of statistical importance. Therefore, I have to make a judgment about the importance of factors and how many of the extracted factors to retain based on various statistical and subjective criteria often found in the literature.

Three statistical methods are most often suggested: Kaiser's K1 rule, parallel analysis, and a scree test (Costello and Osborne, 2005). Each of these examines the eigenvalues that accompany each factor extracted. The eigenvalue is a measure of the amount of variance the measured variables within a factor explains (Beavers *et al*., 2013). For this analysis, I favour the use of parallel analysis, which compares the eigenvalues extracted from the research data to the eigenvalues of a random matrix prior to factor rotation (Franklin *et al.*, 1995). With parallel analysis, factors from the research data that are greater than the eigenvalues of the parallel analysis are recommended to be retained. Parallel analysis is favoured over the Kaiser K1 criterion because it was developed to correct some of the issues of the K1 criterion (Costello and Osborne, 2005; Ledesma and Valero-Mora, 2007). The K1 criterion suggests that the factors retained are those with an eigenvalue greater than 1. This method is noted to be extremely problematic and flawed (Costello and Osborne, 2005; Hoyle and Duvall, 2004; Ledesma and Valero-Mora, 2007; Patil *et al.*, 2008; Yang and Xia, 2015). Even with its

[38] I conducted three tests because the Shapiro-Wilk and the Shapiro-Francia are the most popular, but are common for use with sample sizes of 5000 or less. As each test produced the same significant result of non-normality, I could be confident in the choice of using the PAF.

widespread use, the K1 rule leads to arbitrary decisions because the factors with eigenvalues of 1.01 are considered important compared to factors with eigenvalues of 0.9 (Ledesma and Valero-Mora, 2007). Parallel analysis is also preferred over the scree test. The scree test involves examining the plot of the eigenvalues for breaks in the plots (Hayton *et al.*, 2004). Using this method, the researcher is advised to retain all the factors that come before this break. The justification is that a few factors account for most of the variance in the phenomena of interest and all the factors that come after this break are minor factors. This method can also be problematic because there may not be a clear break in the plot and there may be several breaks. Because of the limitations of the K1 and previous inability to quickly calculate a parallel analysis in statistical software packages, the literature has recommended the use of the scree test (for instance, Costello and Osborne, 2005). This is no longer the case, as Stata 15 is able to produce a parallel analysis via the '*fapara*' command, making obsolete any past concerns about the use of this method.

***Rotation method***

Once the factors have been extracted, they can be challenging to interpret. To facilitate interpretation, the factors usually are rotated to a simple structure. The factor rotation can be either orthogonal or oblique. Orthogonal rotations are the most common choice as they produce factors that are easy to interpret because they are not correlated (Schmitt and Sass, 2011). Oblique rotations, on the other hand, do allow the factors to correlate. Because the factors produced using an orthogonal rotation are not correlated, the subsequent factor structure may not represent reality. This is a noted disadvantage of the orthogonal rotation (Costello and Osborne, 2005). Because most phenomena studied in the social sciences are interrelated and the use of an orthogonal rotation could possibly result in the loss of valuable information, an oblique rotation is adopted as suggested by Matsunaga (2010) and Costello and Osborne (2005). A Promax oblique rotation is used here because it can work well with more massive data sets (Yong and Pearce, 2013).

***Factor loadings & communalities***

Two related features from factor analysis that give us information about the extracted factors are factor loadings and communalities. Once the factors have been extracted, they 'load' onto a factor. Factor loadings give the researcher an idea about how much each variable has contributed to a factor, with larger factor loadings contributing more to the factor (Yong and Pearce, 2013). The factor loadings signify the strength of the correlation between the variable and the factor (Floyd and Widaman, 1995). A factor can be identified by the largest factor loadings. However, I will also examine the small loading or zero loadings in order to confirm the proper identification of the various factors (Yong and Pearce, 2013).

In factor analysis, a simple structure is sought. A simple structure is achieved when each variable loads highly on as few factors as possible, with each variable loading primarily on one factor (Costello and Osborne, 2005; Floyd and Widaman, 1995). There are recommendations I follow from the literature to achieve simple structure. Firstly, I have to determine which cut-off to use for a statistically meaningful rotated factor loading. Though this recommendation can be lowered for data sets with large samples, I will follow a consistent rule of thumb in the literature to retain rotated factor loadings of at least 0.32 (Costello and Osborne, 2005). A 0.32 factor loading provides just over 10% of the overlapping variance (Yong and Pearce, 2013).[39] This factor loading, at an alpha level of 0.01 for a two-tailed test, is considered statistically meaningful (Yong and Pearce, 2013). Secondly, I follow the recommendation that there should be no cross-loading, a situation in which a variable loads higher than 0.32 on two or more factors (Costello and Osborne, 2005; Yong and Pearce, 2013). Avoiding cross-loading will ensure that each factor has a diverse group of interrelated variables and allows for precise, substantive interpretation of the factor(s). Lastly, I also follow Costello and Osborne's (2005) suggestion to keep no factors with fewer than three variables. Fewer than three variables within a

[39] The overlapping variance is equalled to the factoring loading squared.

factor (particularly for smaller data sets) signify a weak factor. In combination, these recommendations should allow for a production of the cleanest factor structure.

Related to the factor loadings are communalities. Communalities are equal to the square of the factor loadings and are the variance in the observed variables that are accounted for by a common factor (Yong and Pearce, 2013). Some recommendations in the literature note that the communalities should be above 0.2 so that there is 80% unique variance (Yong and Pearce, 2013) or less than 0.4 (Costello and Osborne, 2005). Similar to the recommendation for the factor loadings, it has been recognised that this rule can be eased when the sample size is large. For instance, MacCallum *et al.* (2001, 1999) have acknowledged that it is possible to recover factors when communalities are low if the sample size is extremely large. The size of the ACS PUMS can be considered extremely large. I will heed the recommendations of the literature but am also aware that the communalities may be lower than the recommendations because of the large data set. Together communalities and the factor loadings give us an idea about the amount of variance in the factor that is accounted for by each of the variables.

***Checks after analysis***

When theoretically defined variables are used to form a scale, as it happens in factor analysis, they should have internal consistency (Bland and Altman, 1997). Tavakol and Dennick (2011) explain that internal consistency describes the extent to which the variables in a test (factor, in this case) measure the same construct or constructs. Therefore, it is connected to the inter-relatedness of the variables within a factor. The objective test most often used to evaluate internal consistency is Cronbach's alpha (Cronbach, 1951). Cronbach's alpha is used to estimate the proportion of the variance that is consistent in a factor. If the items in a factor are correlated to each other, the value of alpha is increased. The value of alpha can range from 0 to 1, with 0 suggesting that no variance is consistent and 1 meaning that there is perfect variance consistency. Cronbach (1951) did not provide a cut-off for alpha, but others have later noted that anything below 0.5 is an unacceptable value (George and

Mallery, 2003). I recognise, however, that a high coefficient for alpha does not necessarily mean a high degree of internal consistency (Schmitt, 1996; Tavakol and Dennick, 2011) because alpha is also affected by the number of variables that make up the factor. If there are few variables within a factor, the value of the alpha will also be reduced. Recall that I will follow a recommendation from Costello and Osborne (2005) to have at least three variables per retained rotated factor. Having a factor with three variables may result in a low value of alpha. Addressing this limitation of the alpha, Schmitt (1996) suggests not to report alpha alone, stating that "presenting only alpha … is not sufficient. Intercorrelations … must be presented as well" (p. 353). In addition to presenting the Cronbach's alpha to measure internal consistency, data will be presented regarding the intercorrelations for each factor.

***Factor scores***

Factor analysis plays a key methodological and substantive role in this thesis. The information derived from the factor analysis will be used in subsequent analysis. In order to use the EFA information in a follow-up analysis, factor scores have to be calculated.[40] Factor scores represent each sample member's placement on the factor(s) identified in the EFA (DiStefano *et al.*, 2009). Factor scores can be determined using non-refined and refined methods. I adopt a refined method in this analysis because these methods aim to maximize validity by producing scores that are correlated with a given factor (DiStefano *et al.*, 2009). Two refined methods are available: regression factor scores and Bartlett factor scores. Regression factor scores are not suitable when the factor scores are used as dependent variables because the regression parameter will be biased in most cases (Devlieger *et al.*, 2016). As the factor scores here will be used as dependent variables only in latter analyses to represent the multiple dimensions of disadvantage in the United States, the Bartlett

[40] Structural equation modelling is an option here. It is not used in this thesis because it is best used in conjunction with CFA (Gallagher and Brown, 2013). The goal of linear structural equation modelling is to generalise the CFA model in order to assess how the derived variables are related to other variables (Lee and Song, 2010). As acknowledged in section 6.2.2, I utilise EFA in this analysis. I proceed with deriving the factors scores as outlined in this section.

factor scores are used. The resulting products are continuous variables that are calculated as Z scores. They have an average of zero and a standard deviation close or equal to zero. A sample member's placement on each factor score will give some indication of how disadvantaged or advantaged they are on each factor of social exclusion.

The resultant factors from the EFA help to answer the primary research questions of this thesis:

1. What is/are the factor(s) of multidimensional disadvantage in the United States? (Chapter 7)
2. To what extent are sociodemographic characteristics associated with multidimensional disadvantage in the United States? (Chapter 8)
3. Is there variation in multidimensional disadvantage across the United States and the District of Columbia? (Chapter 9)

The first research question depends heavily on the results of the EFA. I will be able to refute or lend weight to an argument that disadvantage in the United States is multidimensional. The second research question follows a long history in the United States and examines the relationships between individual sociodemographic characteristics and disadvantage as I highlighted in Chapter 2. For the third research question, the analysis takes advantage of the nested hierarchical nature of the data and explores individual disadvantage differences between states. Sections 6.3 and 6.4 address the methodological techniques utilised to answer research question 2 (OLS multivariate linear regression) and 3 (multilevel modelling), respectively.

## 6.3 Testing the association between disadvantage and individual characteristics using multivariate regression

The second methodological objective is to examine the association between individual sociodemographic characteristics and multidimensional disadvantage in the United States. In this case, I am concerned if there is a statistical relationship between response variables - the factor(s) of disadvantage - and independent variables (the sociodemographic characteristics). I employ an Ordinary Least Square (OLS) multivariate linear regression analysis. I choose this statistical method because OLS multivariate regression is a technique that allows more than one independent variable to be present in the models, allowing for the concurrent examination of the relationships between the independent variables in the analysis and whilst controlling for others. This will provide insights into whether different groups experience disadvantage differently. In addition, linear regression requires a continuous dependent variable. The final construct(s) derived from the factor analysis is/are continuous, conceptually distinct variables that approximate a normal distribution, making the OLS linear regression suitable for use with the factors of multidimensional disadvantage (Cleophas and Zwinderman, 2018; Hutcheson, 1999). The equation of the OLS multivariate regression model is as follows:

**Equation 6.1: Multivariate regression equation**

$$y = \alpha + \beta_1 x_1 + \beta_2 x_2 + \ldots + \varepsilon$$

Where:

y = dependent variable

α = constant

β = coefficient

x = independent variable

ε = error term

The regression model predicts the dependent variable (each factor of multidimensional disadvantage) based on X. The various X's are the individual sociodemographic characteristics. These include race, gender, age, citizenship status.

A special note in this analysis is that each of the independent variables, except age, are qualitative, nominal variables. In order to include these into the regression analysis, they are transformed into dummy variables in order to represent group membership. Dummy variables are subject to a type of scoring called dummy coding, which involves transforming a qualitative (also referred to as categorical) variable with *n* categories into *n*-1 dummy variables (Bech and Gyrd-Hansen, 2005; Jaccard and Turrisi, 2003). It involves assigning one (1) to all members of one group and zero (0) to everyone else. Hutcheson (1999) acknowledges that OLS regression is a powerful technique for modelling continuous data in conjunction with dummy variable coding and data transformation. Therefore, I am confident that this technique will allow for adequate testing of the association between and within subgroups and their relationship to multidimensional disadvantage.

I acknowledged in Chapter 2 the importance of adopting an intersectional framework in the analysis of disadvantage in this thesis (section 2.4.1). In section 5.3, I acknowledged that an intercategorical method for assessing intersectionality quantitatively is adopted in this thesis, in which I use existing analytical categories to understand disadvantage in the United States. From here, I have two options to measure intersectionality using McCall's (2005) intercategorical approach: the additive and the multiplicative approach (Dubrow, 2008). Both can be achieved using an OLS multiple regression model, but I choose to use the multiplicative approach. This is chosen for two key reasons. Firstly, intersectionality, at its core, emphasises that the influence of a given demographic characteristic on a social outcome is conditional on the intersection of characteristics (Dubrow, 2008). In the multiplicative approach, this approach begins with the characteristics that make up the intersections (Dubrow, 2008; McCall, 2005). Then the characteristics that make up the intersection, which have valid meaning and provide context, are examined. The additive approach misses this core premise of intersectionality. It assumes that each characteristic of an individual has an additive effect and should be measured separately (Dubrow, 2008). In an intersectional analysis, these characteristics are always experienced in tandem within an individual (Johnson and Loscocco, 2015). Secondly, intersectionality via the multiplicative approach can be examined by adding an interaction term in the multivariate regression. Bowleg (2008) acknowledges that interactions between constructs are at the heart of intersectionality research. Adding an interaction term into the OLS regression can allow for an examination of how the intersection between race and gender, the social categories I am interested in for this thesis, [41] confers a unique experience above and beyond being a member of either group. These regression models are often called the OLS moderated multiple regression (Disatnik and Sivan, 2016). The equation for this is as follows:

**Equation 6.2: Multivariate regression equation with interaction terms**

$$y = \alpha + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1 x_2 \ldots + \varepsilon$$

Where:
y = dependent variable
α = constant
β = coefficient
x = independent variable
ε = error term

The term, '$\beta_3 x_1 x_2$', in Equation 6.2 represents the interaction effect. In order to determine the statistical significance of the interaction, it is necessary to include the two independent variables that make up the interaction in addition to the interaction term. Therefore, it is imperative to keep the

[41] The intersection between race and gender was the primary intersection of concern for Crenshaw (1991) in articulating the concept. I retain that focus.

variables for race and gender in the model as well. The other components of Equation 6.2 are the same as 6.1.

### 6.3.1 Interpreting the OLS multivariate regression model

When interpreting the regression model, it is best to consider the importance of the independent variables and the types of relationships found. The coefficients, which can either be unstandardized or standardised, gives us some indication of the relationship between the independent and dependent variable, whether it is positive or negative, strong or weak. The unstandardized coefficient $\beta_1$ can be interpreted as the change in the mean of the distribution in the dependent variable associated with a unit change in X, holding the other variables in the model constant (Craven and Islam, 2011; Jaccard and Turrisi, 2003; Montgomery *et al*., 2012). Standardised coefficients do not depend on the units of the independent variable. The standardised coefficient refers to how many standard deviations the dependent variable will change per standard deviation increase in the independent variable (Nieminen *et al.*, 2013). These coefficients cannot be generalised to the wider population as they have been standardised according to the scales of the sample data. Subsequently, I will not be able to compare these effect sizes. Instead, I will rely on the unstandardized coefficients for the interpretation of the coefficient size and to ensure generalisability to the wider population in the United States.

Montgomery *et al.* (2012) acknowledge addressing the adequacy of the regression model. Therefore, two goodness of fit measures will be used to interpret the overall fit of the model: 1) R-square (Jaccard and Turrisi, 2003) and 2) the F-statistic (Hutcheson, 1999). The R-squared ranges from zero (indicating no linear relationship) to one (indicating a perfect linear relationship). In order to give the percentage of variability explained by the model, it will be multiplied by 100. The F-statistic provides a measure of the significance of the model (Hutcheson, 1999). If the value of the F-statistic is not significant, a hypothesis of no relationship between the dependent and independent variables is accepted. Contrarily, if the value of the F-statistic is significant, we can accept that there is a relationship between the variables. Because of these checks, the regression model is an iterative process, in which the data lead to a model and the fit of the model leads back to the data (Montgomery *et al.*, 2012).

## 6.3.2 Weaknesses of the OLS models

In order to utilise the OLS multiple regression method, the researcher is bound by a strict set of assumptions regulating both the dependent and independent variables. Below is a discussion of these assumptions as is relevant to this analysis.

There is an assumption of normality of the dependent variable. Section 6.2.2 acknowledged that the constructs derived from the factor analysis will be used as the dependent variable(s) in the OLS model. These constructs are presented as having a continuous level of measurement. We can assume that these variables are normally distributed, but I recognise that particularly due to the component variables of the factor analysis, these dependent variables may not be perfectly normally distributed. However, issues of extreme diversions from normality are noted to be an issue in small samples (Lumley *et al*., 2002). Lumley *et al.* (2002) recognise that in large samples, the use of the linear regression is useful for any distribution of the variable. Therefore, issues of abnormality (or extreme abnormality), if they arise, should not be an issue for this analysis.

There is an assumption of independence, meaning that data from different respondents are independent. As discussed in Chapter 4, the United States' Census Bureau utilises a sophisticated sample design that obtains information from specific addresses once every five years (section 4.2.1). As this analysis utilises just one year of the ACS PUMS data set, we can be sure that the same addresses are not contacted for survey response during the year of analysis. On the other hand, we cannot be sure that there is no independence within each state or county. For just one year of ACS PUMS data, areas with a population of over 65,000 residents are included. In these areas, it is difficult to say that the respondents at those addresses are independent. It can be assumed that people within these geographical areas have some similarities when compared to a completely random sample. To overcome this weakness, robust standard errors can be calculated, taking into account the complex sample of the ACS PUMS, via the use of the *surveyset* command in Stata 15. With the *surveyset* command, I can apply the PUMS' eighty replicate weights and make corrections for the sample. This technique is recommended by the Census Bureau (2014) for use as well.

The OLS regression assumes that multicollinearity is not present. Multicollinearity can occur when the independent variables are highly correlated, producing inflated standard errors of the regression coefficient (Disatnik and Sivan, 2016). For the regression model presented in Equation 6.1, sociodemographic characteristics are entered as the independent variables. I am expecting that there will be no multicollinearity issues present with these variables, but I will conduct the appropriate test to ensure it is not violated. In the moderated multiple regression model, Jaccard and Turrisi (2003) and Disatnik and Sivan (2016) argue that multicollinearity between a product term and the component variables will generally not be challenging for interaction analysis. There are no other expected violations of the regression model.

In light of the weaknesses and strict requirements of the OLS models, I considered a partial least squares model in order to answer the second research question. The benefit of partial least squares is that it relaxes some of the assumptions necessary for an OLS regression analysis. For instance, the partial least squares method can handle many independent variables even if there is a presence of multicollinearity and can work with small sample sizes (Pirouz, 2006). This method was not selected for use here because the advantages that it has over OLS is not necessary for this analysis. There will not be a multicollinearity issue as the sociodemographic characteristics are discrete categories. In addition, the sample size utilised for this analysis is sufficiently large. Lastly, with partial least squares, there is a lack of model test statistics that are present with OLS that can aid in determining the statistical significance of the models used. Subsequently, it was decided that the OLS regression was the optimal method as the noted disadvantages of the method are not considered to be problematic with the data used.

The use of the OLS multivariate regression is chosen to analyse the relationship between individual sociodemographic characteristics and the factor(s) of social exclusion in the United States. It is the most appropriate technique to reach this aim of the thesis for methodological and substantive concerns. In the first instance, OLS regression modelling is iterative in which the analysis begins and ends with the data. In the second instance, with multivariate linear regression modelling, I am able to consider the relationship between many independent variables and the dependent variable(s), holding the other variables in the model constant. This offers an excellent introduction to exploring multidimensional disadvantage in the United States and exploring its relationship to individual characteristics.

## 6.4 Exploring state-level variation in disadvantage

The third methodological and substantive objective lies with determining whether any variation exists in multidimensional disadvantage across the United States and the District of Columbia. The most appropriate method to achieve this objective is multilevel modelling (or linear mixed modelling). This is a suitable statistical technique for reasons related to the structure of the ACS PUMS data and the associated research question (**Is there variation in multidimensional disadvantage across the United States?**). Firstly, multilevel modelling is geared toward the analysis of complex data structures that are hierarchical or clustered in nature (Harrison *et al.*, 2018; Hox, 1998; Peng and Lu, 2012). The ACS PUMS fits this structure as it clusters individuals in households and households in each state. These types of data often require more advanced models. Secondly, multilevel modelling is a statistical technique for large samples. Hayes (2006) notes that the mathematics and the theoretical assumptions that underlie the statistics of multilevel models are based on the behaviour of statistics calculated in large samples. This is important in terms of the validity of the inferential tests as well as their ability to detect a relationship worthy of detection. With a data set that contains 2,348,374 sample members, this criterion for multilevel modelling is more than met. Next, the respective research question adds the state as an additional level of analysis beyond the individual. There is specific interest in how the grouping of individuals into states is associated with each outcome of disadvantage, which is well suited for multilevel analysis (Dieleman and Templin, 2014). The research question is important because nested datasets do not automatically require a multilevel model (Peugh, 2010). If variation is found between the states, then an MLM is needed to estimate the variance in multidimensional disadvantage that occurs both across individuals and across states. Traditional multivariate regression techniques can only model the variance in disadvantage at a single unit of analysis (the individual ***or*** the state). Finally, a linear version of the multilevel model is chosen because the factor analysis produces continuous measures that are used as the dependent variables. While multilevel modelling has been used as methodological technique in poverty-related studies (Brady *et al.*, 2013; Chen and Wang, 2015; Kim *et al.*, 2010; Subramanian *et al.*, 2005), most conceive poverty as a binary outcome as income poor vs. not income poor. Subsequently, these use a multilevel logistic model. In Chapter 3, I argued against constructing social exclusion as a binary outcome, so a multilevel logistic model is not appropriate here.

At its core, the MLM is not very different from single level regression models (as discussed in section 6.3). In the MLM, the outcome variable - the factor(s) of disadvantage - is modelled as a linear combination of independent variables that are each weighted by a coefficient that quantifies the relationship between that independent variable and the outcome (Hayes, 2006). The primary difference between the MLM and single-level models is the MLM's ability to estimate one or more of the coefficients in the model as either fixed or random (Harrison *et al.*, 2018; Hayes, 2006). The ability to estimate random effects has several benefits. Firstly, the data set is structured in such a way that using a single level model to answer the research questions would violate that model's assumption of independence among the observations (Dieleman and Templin, 2014). The observations in the ACS PUMS data are collected from individuals within each state. Subsequently, we might expect that the measurements within a statistical unit (here individuals in states) may be more alike than measurements from different states. Multilevel modelling allows us to explicitly model this non-independence in hierarchical data. Modelling the random effects will ensure correct inference about the fixed effects (Dieleman and Templin, 2014; Harrison *et al*., 2018). In addition, if fixed effects vary at the level of the individual, then non-independence within states could also be accounted for. The random effects in this analysis represent the grouping variable, the state, and allows us to estimate the variance in the disadvantage outcome(s) within and among states (Harrison *et al.*, 2018). This subsequently reduces the probability of false positives (Type I errors) and false negatives (Type II errors) (Harrison *et al*., 2018; Huang, 2018). Being able to infer the magnitude of variation within and among states of multiple forms of disadvantage is a substantial methodological and substantive contribution of this thesis.

### 6.4.1 Fitting the models

In this section, I discuss the approach for each of the related research questions. It was recognised in Chapter 5 (section 5.4) that there is insufficient data for the various indicators of social exclusion for individuals under the age of 18 and the group quarters populations. Once those observations are excluded from the analysis, the number of individuals per household averages at one, negating the differential experiences within the household. Subsequently, the analysis does not nest individuals within households and households within states. Individuals are nested within states, [42] supporting a two-level model.

There are two specific decisions made before conducting the multilevel analysis that has to be addressed: the estimation method and the use of weights. Firstly, the researcher has to decide on the estimation method (Hayes, 2006; Peugh, 2010). There are two estimation methods offered in Stata 15 for multilevel modelling: maximum likelihood and restricted maximum likelihood. This thesis opts for the maximum likelihood method as most of the reasoning offered to use restricted maximum

[42] I follow other studies (e.g. Brady *et al*., 2013) and classify the District of Columbia as the 51st state. This is done purely for ease of analysis. As recognised in section 2.5, DC is technically not a state, but fulfils the same roles as official states for the purpose of this analysis, particularly in understanding the role of context.

likelihood centres around its ability to provide more accurate variance estimates when the sample size is small (McNeish, 2017). As the data utilised in this thesis have a substantially large sample size, particularly at the individual level (where n > 2 million), these corrections are not needed.[43] Recognising that the level 2, state-level, units are always smaller than the individual level units, Maas and Hox (2005) suggest that when the number of units at level 2 are 50 or less, there is potential for biased estimates of the standard errors at level 2. For this thesis, there are 51 level 2 units, which is the entire population of states in the United States. The availability of data for all possible level 2 units suggests that the bias produced by a small number of level 2 units is avoided in this analysis.

Secondly, I make the decision not to use any weights, be it analytical or survey, when estimating the multilevel models. Carle (2009) found that when fitting multilevel models to complex survey data, there were minimal observed differences between weighted and unweighted analysis. Additionally, with the sample size over 40,000, Carle's (2009) experiment did not lead to different inferential conclusions. With the large sample size of the ACS PUMS data, the use of weights is likely to not yield substantially different results, compared to a non-weighted analysis.

### *6.4.1a Exploring variation in multidimensional disadvantage across the United States*

The most basic multilevel model is the variance components or null model. The null model is also called an 'empty model' because it has no predictor variables (Hayes, 2006). This model will be used in this thesis because it focuses on assessing whether the state level (level 2) units differ from each other on the outcome variables (the resultant latent constructs from the factor analysis). The multilevel equations for this model are:

[43] Statistics certainly may be affected by a certain size of the data. A small sample size is often recognised to be problematic in many different instances, as noted throughout this chapter. Utilising data with such a large sample mean that most of the requirements for certain statistical techniques are more than met. However, I recognise that such a large sample can be problematic in its own right, particularly if the data are not representative of the population to which I am generalising the results (Kaplan *et al.*, 2014). Additionally, there is a potential that statistical significance is much more likely to be achieved with such a large sample. In Chapter 4, I addressed the procedures taken by the Census Bureau regarding representativeness. Lastly, the results in Chapter 8 will show that not all the relationships tested using this data and the methods discussed in section 6.3 were statistically significant, even with such a large sample.

**Equation 6.3: Multilevel equation form of the unconditional model**

$$\text{Level 1: } Y_{ij} = \beta_{oj} + r_{ij}$$
$$\text{Level 2: } \beta_{oJ} = \gamma_{oo} + u_{oj}$$

$Y_{ij}$ is the $i^{th}$ individual score on the outcome variable in state *j*. $\beta_{oj}$ is the average score on the disadvantage factor for all individuals within state *j*. $r_{ij}$ is a residual term that quantifies the difference between the state mean score on the factor(s) and an individual's score on the factors. In the Level 2 model, $\gamma_{oo}$ is the average score on the respective factor of disadvantage; it is essentially a grand mean aggregating the averages across states. Lastly, $u_{oj}$ is the difference between state j's average and this grand mean.

This equation is also presented in a mixed form.

**Equation 6.4: Mixed model equation form of the unconditional model**

$$Y_{ij} = \gamma_{oo} + u_{oj} + r_{ij}$$

This form of the model shows that an individual score in a disadvantage factor in state *j* is a function of three different components: 1) the average disadvantage score in each state ($\gamma_{oo}$), 2) how much state *j*'s disadvantage score differences from the grand mean ($u_{oj}$), and 3) and the difference between individual *i*'s score and the state *j*'s average score. This model is known as a random intercept model, as the measure $\gamma_{oo}$ + $u_{oj}$ is the random intercept containing a fixed component $\gamma_{oo}$ and a random component $u_{oj}$ (Hayes, 2006). $r_{ij}$ is also a random component.

Multilevel models estimate the variances of the residuals at level 1 and level 2, not the actual residuals (Hayes, 2006; Peugh, 2010). This means that the variation in disadvantage scores at level 2 quantifies the variation in average disadvantage scores across states (Peugh, 2010). Peugh (2010) suggests two statistics to determine how much variation is present among the states (level 2 units):

the intraclass correlation coefficient (ICC) and the design effect. The ICC, which is conceptually like the R-square of OLS regressions, represents the amount of variance attributable to the state and is estimated using the unconditional model that has no independent variables (Huang, 2018; Peugh, 2010). Kreft and de Leeuw (1998) further acknowledge that the ICC is a measure of the degree of dependence of individuals and can be summarised in a number of different ways. For the purposes of its uses here, the ICC indicates the proportion of the variance in multidimensional disadvantage that is between each state. Peugh (2010) suggests that an ICC value of zero indicates two things: 1) there is no variation in disadvantage across states and 2) all variation in disadvantage score occurs across individuals. The latter point would indicate that the relationship between disadvantage and individual sociodemographic characteristics is, on average, the same across states.

The design effect measures the effect of independence violations on standard error estimates (Peugh, 2010). With nested data, there may be homogeneity within each state that would lead to underestimated standard errors. Thomas *et al*. (2005) acknowledge that the main idea is that the more similar observations are within their respective clusters (states), the greater the likelihood for underestimating the true variability of the population. So the design effect is an estimate of the multiplier that has to be applied to standard errors in order to correct for this type of bias that is a consequence of nested data (Hox, 1998; Peugh, 2010). Peugh (2010) recommends that the design effect should be greater than 2, as anything lower than that would suggest a multilevel model is not needed. Lai and Kwok (2015) acknowledge that 2 is an often used cut-off for the design effect, but after a Monte Carlo simulation study found that this rule should not apply in two cases: 1) when the researcher is interested in the effects of higher-level predictors and 2) when the cluster size is less than 10. This thesis does not meet any of these criteria, as state-level (level 2) predictors are not included in this analysis and the cluster size is 51. Therefore, I will apply Peugh's (2010) suggested cut-off to this analysis.

In addition to these statistics, Hayes (2006) notes that evidence of variance can be determined by assessing whether the variance of the random parts of the intercept, $u_{0j}$, is different from zero (Hayes, 2006). If the states are only slightly different, then the *j* values of $u_{0j}$ should differ little from each other and exhibit little to no variance.

If the null models suggest that there is significant variation in disadvantage across the states, then it becomes possible to add additional components to the MLMs. Therefore, contingent upon the results of the null model, I have selected two additional research questions that can also be answered via the use of multilevel models. The questions as well as the approach to answer those questions, if needed, follow.

### *6.4.1b Does that variation, if any, still persist after controlling for individual characteristics?*

Should the null models suggest significant variation exists across the states in disadvantage, multilevel modelling can be used to assess if that variation persists whilst controlling for individual characteristics, particularly age, race, gender, and citizenship status. In this case, I will employ random intercept multilevel models, one for each factor of social exclusion uncovered from the EFA.

In building a random intercepts MLM, there is an additional choice made that influences the specification of the model at level 1: the centring of level 1 variables (Peugh, 2010). Centring essentially involves rescaling an independent variable so that a value of zero can be interpreted meaningfully (Peugh, 2010). All of the independent variables of interest in this study are sociodemographic characteristics which are added in the models as dummy variables. As dummy variables have a meaningful zero, I follow a recommendation by Nezlek (2012) not to centre dummy variables at level 1. Subsequently, the analysis can proceed without further consideration of centring.

The equation for the random intercepts model in multilevel form is as follows:

**Equation 6.5: Multilevel equations for the random intercepts multilevel models**

$$\text{Level 1: } Y_{ij} = \beta_{0j} + \beta_{1j} + X_{ij} + r_{ij}$$
$$\text{Level 2: } \beta_{0j} = \gamma_{00} + u_{0j}\ ;\ \beta_{0j} = \gamma_{10}$$

$Y_{ij}$ measures the relationship between the outcome variable as a function of the respective characteristics of the individual. Note that similar to Equation 6.1, the $\beta_{1j}$ signifies the coefficient for the respective variable X in the model. In fact, Hayes (2006) acknowledges that $\beta_{1j}$ is conceptually equivalent to the unstandardized regression coefficients in a single level regression and can be interpreted as such. The interpretation of the unstandardized regression coefficients is addressed in section 6.3.1

The mixed form of Equation 6.5 is as follows:

**Equation 6.6: Mixed form equation of the random intercepts multilevel model**

$$Y_{ij} = \gamma_{00} + \gamma_{10} X_{ij} + \gamma_{20} X_{ij} + \ldots + u_{0j} + r_{ij}$$

Here, $\gamma_{10}$ and $y_{20}$ measure the relationship between the sociodemographic characteristic (X) and disadvantage. This relationship is fixed and remains constant across states. This means that there is only one coefficient estimating the effect of that variable (Hayes, 2006). For instance, if X is a dummy variable for gender (female), then the respective coefficient represents the average difference between men and women across states. This assumption of fixing the coefficients to a constant value is an assumption that can be tested, which will be discussed in section 6.4.1c.

After adding the predictors to the conditional model, it is possible to test if the fit of the model is improved. There are two tests that I will use: likelihood ratio test and the deviance (Hayes, 2006; Peugh, 2010). Consider that the null model is nested within the random intercepts model. The likelihood ratio test is a statistical test of two nested models (the null model and the random intercepts model, in this case) (Peugh, 2010). If the test is statistically significant, we can be confident that the addition of these predictors improves our understanding of disadvantage.[44] In Stata 15, the likelihood ratio test is performed during the calculation of the MLM, rendering it relatively easy to determine the fit of the model.

The deviance is used to test the hypothesis that additional model predictors do not improve the fit of the model (Peugh, 2010). The maximum likelihood estimation method uses log-likelihood to measure the probability that the model being estimated produced the sample data. Multiplying this value by -2 produces the deviance. In order to test the fit, the deviance of the random intercepts is subtracted from the deviance of the null model. The result is tested in the chi-distribution table with the degrees of freedom equalling the difference in parameter estimates between the models (Hayes, 2006). In this case, the only difference between the models would be the predicted value of the coefficients (for instance, $\gamma_{10}$ and $\gamma_{20}$). A significant result of this indicates that including the predictors are a better fit for the model than having no predictors. The ICC and design effect can also be calculated for the random intercepts models. Calculating these will have the same role as they do for the null models.

#### *6.4.1c Does the relationship between individual sociodemographic characteristics and multidimensional disadvantage vary significantly across states?*

The models discussed in section 6.4.1a and 6.4.1b fix the relationship between the independent variables and the factor(s) of disadvantage. It could be that states differ in the level of disadvantage with respect to various individual characteristics, for instance, because of the proportion of people in the states with different characteristics. MLM allows us to investigate whether the effect of a level 1 variable varies across level 2 units without actually specifying what those differences might be.[45] We

---

[44] This test can also be conducted to compare the null model to a single level model. A significant result indicates that a multilevel model is a better model fit than that of a single level.

[45] Because state level predictors are not included in the ACS PUMS, state level predictors are not included in the analysis. Combining multiple data sets are beyond the scope of this analysis. Via the model identified in section 6.4.1c, I am able to assess if there are state level differences without hypothesising and testing them here. In addition, the relevant research question is only concerned with examining if the relationships between the independent variables and the disadvantage factor(s) vary across the state.

can do this via the use of a random coefficient MLM.[46] In this model, the effect of level 1 variables are set as random and then the variance of the random component is assessed to see if it is statistically different from zero (Hayes, 2006). Essentially, I will extend the model discussed in section 6.4.1b by estimating the coefficient for the independent variables as a random effect. The multilevel form of the random coefficient MLM is as follows:

**Equation 6.7: Multilevel equation random coefficient multilevel models**

$$\text{Level 1: } Y_{ij} = \beta_{0j} + \beta_{1j} + X_{ij} + \beta_{2j} + X_{ij} \ldots + r_{ij}$$
$$\text{Level 2: } \beta_{0j} = \gamma_{00} + u_{0j;}\ \beta_{1j} = \gamma_{10} + u_{1j;}\ \beta_{2j} = \gamma_{20} + u_{2j}$$

The mixed form of the model is as follows:

**Equation 6.8: Mixed equation for the random coefficient multilevel model**

$$Y_{ij} = \gamma_{00} + (\gamma_{10} + u_{1j})\,(X_{ij}) + \gamma_{20} + (X_{ij}) + \ldots + u_{0j} + r_{ij}$$

The $u_{1j}$ in the level 2 model of $\beta_{1j}$ represents the random part of the effect of the independent variable (Hayes, 2006). This is what allows the independent variable(s) to vary across states. $\gamma_{10}$ then is the average effect the independent variable across states and the $u_{1j}$ measures how the effect of the independent variable for state j differs from the average (Hayes, 2006).

[46] These models are also called random slopes. When the independent variables that are to be set to random, (rather than fixed) are categorical (rather than continuous), the terminology changes slightly to random coefficients. Though they have been used interchangeably across the literature, I will refer to them as random coefficients because each of the independent variables in this analysis are categorical.

In the random coefficient model, there is more than one random effect. When this happens, it becomes possible to measure their individual variances as well as their relationship with each other (their covariance) (Hayes, 2006; Peugh, 2010). In the model, $u_{1j}$ and $u_{0j}$ are both random effects. $u_{0j}$ denotes the random component of the intercept for state j (the difference between the intercept for state j and the average intercept across states); $u_{1j}$ denotes the random effect of an independent variable for state j (the difference between the effect of the independent variable for state *j* and the average effect of the independent variable) (Hayes, 2006). When estimating the covariance, I will opt for the unstructured covariance type, as opposed to a variance components type. It will allow these random components of the random coefficients MLM to intercorrelate reflecting the fact that the states with relatively more or less disadvantage than another state may be more or less affected by the independent variables in the model (Hayes, 2006). In addition, the unconstructed covariance type has been suggested for use elsewhere in the literature because it does not conform to any systematic patterns (Field and Wright, 2011; Gurka *et al.*, 2011). In addition, Hayes (2006) acknowledges that forcing the random components not to correlate would be "unnatural in most circumstances" (pg.401).Given the expected interrelated nature of social exclusion and the domains of the B-SEM, it would be remiss not to allow correlation.

The interpretation of the covariance can be facilitated by converting it to correlation, which is accomplished by dividing the covariance by the square root of the product of the variance of the random intercept components and the variance of the random components of the independent variable (Hayes, 2006). Stata 15 calculates the correlation via the '*estat recovariance, correlation'* command. Hayes (2006) does acknowledge the significance test associated with the random effects correlation may conflict with the likelihood ratio test, so he and Peugh (2010) suggest that the likelihood ratio test is the better test of the random effects. Where able, I will attempt an interpretation of the correlation, but I determine the statistical significance of the random effects by the likelihood ratio test (Hayes, 2006). With the likelihood ratio test, I will be testing if the random coefficient model is preferred over the random intercepts. It makes sense not to compare the random coefficients MLMs to the null model because it is an extension of the random intercept model.

### 6.4.2 Effect size of multilevel models

In an OLS regression model, we can test the effect size using the R-square. As acknowledged in 6.3.1, using the R-square, we can determine how much of the variation in the dependent variable can be explained by the independent variables in the model. In a multilevel model, it is not possible to calculate an R-square. Recognising that the variation in a dependent variable in multilevel models are partitioned between level 1 (individual) and level 2 (state) components, a pseudo-R square type statistic can be calculated (Peugh, 2010). This is called the 'variance accounted for' measure (Hayes, 2006). By comparing the estimated residuals between two nested models, the variance accounted for measure tells us how much of the variance remaining in the dependent variable is attributable to the independent variables included in the model (Hayes, 2006). The 'variance accounted for' measure is equalled to 1 – the ratio of the two variance estimates. The result benchmarks variance explained by the independent variable relative to the variance remaining after removing variance accounted for by differences between states in disadvantage (Hayes, 2006).

## 6.5 Conclusion

In order to reach the substantive and methodological aims of this thesis, I have chosen three statistical techniques: 1) exploratory factor analysis, 2) ordinary least squares multivariate regression and 3) multilevel modelling. In order to measure social exclusion, a latent construct, I have concluded that EFA is the most appropriate method to reduce the observed indicators identified in Chapter 5 into the underlying construct(s). To explore the association between individual sociodemographic characteristics with the constructs of multidimensional disadvantage, OLS multiple regression is utilised. OLS multivariate regressions are also used to test whether there is a moderating effect of race on a women's experience of disadvantage via the addition of interaction effects to the OLS models. Lastly, multilevel modelling is used to explore whether there is any variation in disadvantage at the state level. These techniques are subsequently applied to the 2015 ACS PUMS data and inform the final part of the thesis: the results.

# PART 3: THE RESULTS

Part 3 of this thesis comprises of three findings chapters. Chapter 7 uncovers the factors of disadvantage in the United States employing the EFA on American data. Chapter 8 tests the associations between the factor(s) of disadvantage with sociodemographic characteristics, including an application of intersectionality. Chapter 9 presents the results of the multilevel models that consider the state as a level of analysis. In Chapter 9, I examine the role of context in understanding social exclusion in the United States.

After the findings chapters, I conclude the thesis in Chapter 10. There, I discuss the implications of the analysis undertaken within as it relates to policy. I also acknowledge the limitations of the research and its contributions and offer suggestions for areas for future research.

# Chapter 7

# Deriving the factors of (dis)advantage in the United States

## 7.1 Introduction

The main objective of this chapter is to derive the latent construct(s) of social exclusion utilising 16 variables from the American Community Survey (ACS) Public Use Microdata Sample (PUMS), as described in Chapter 5. Exploratory Factor Analysis (EFA), a technique that has been used in human research development (Reio Jr and Shuck, 2015) and psychology (Hirschfeld *et al.*, 2014), was recognised in the previous chapter as the best available method to reduce the theoretically identified variables to capture the nature of social exclusion into interpretable factors. The resultant construct(s) are used to answer the following research question:

*What is/are the factor(s) of multidimensional disadvantage in the United States?*

## 7.2 The challenges of analysing a huge sample

The challenges of utilising a substantially large data set for the analysis have been considerable. With a data set that includes over 2.3 million sample members, there was limited computing capacity to complete a rigorous analytical task like factor analysis. I was able to obtain access to a medium power computer in July 2016,[47] but the sheer size of the data and the amount of time needed to complete

[47] At the School of Social and Political Science at the University of Edinburgh, a medium power computer is a multi-user workstation that students can use to carry out long-running analysis. It is a high specification workstation stored in a computer room at the Chrystal MacMillan Building. Specifications for the medium power computer are as follows:

- HP Z230 Tower Workstation
- Intel (R) Xeon (R) CPU E3-1245 v3 @ 3.40GHz – 4 Core with Multithreading giving 8 Cores
- 2TB SSHD Model ST2000DX001-1CM1 – Solid State Disk Drive
- 32GB RAM

tasks was too great for the medium power computer.[48] In order to combat the issues of processing large amounts of data, a random subsample of 10% was taken. From here, only the random subsample with n=234,846 is used and discussed. The smaller data set still has a large enough sample that will allow for generalisability and meet the criteria in order to conduct the relevant statistical analysis.[49]

Due to continued challenges with the rather large subsample in producing a correlation matrix necessary for factor analysis, I opted for a mixed correlation method for the initial stages of the analysis.[50] I recognised in the previous chapter that a polychoric correlation is advised for my set of variables (section 6.2.1). A polychoric correlation matrix was attempted on the random subsample of the data, but a usable matrix was not obtained utilising Stata 15. When attempted via command '*polychoric,*' a command recognised to be slower than the '*corr*' command for Pearson's correlation, it ran for several weeks without delivering a matrix. During its run, the phrase '*numerical derivatives are missing*' appeared in the output window, highlighting that the matrix had missing values. As Stata cannot complete a factor analysis with missing values in the correlation matrix, it became necessary to find other suitable methods to ensure sufficient correlation among the variables. [51] [52] Therefore, several Cramer's V tests were conducted for the nominal variables and Pearson's correlations for the

---

[48] During an extended session of producing some results, the hard drive of the medium computer power burned. This halted analysis until the hard drive could be replaced.
[49] I am not aware of a maximum sample size to conduct factor analysis. Most literature, some of which are acknowledged in section 7.3, are limited to a discussion on the minimum sample size.
[50] The use of various correlation methods is only done here to ensure that there is some correlation amongst the variables in their original form. As discussed in section 7.3, I am able to create dummy variables for many of the variables included in the factor analysis (presented in Table 7.1) and able to use a more common correlation technique, Pearson's. Pearson's correlations are conducted in Stata in the processing of factor analysis, using the command, *factor.* Therefore, correlation among the variables is tested twice in the analysis.
[51] A tetrachoric matrix was attempted by converting some of the variables to binary, meaning they take the value of 0 and 1. The tetrachoric matrix also did not work in Stata 15. Even if it worked, a tetrachoric correlation would not have been ideal either as it requires a binary variable for all the variables, including the continuous variables, such as 'travel time to work.' It would have been particularly difficult to categorise the travel times into useful binary variables. The same is true for income.
[52] I also attempted to address the problem of missing values. The Expectation Maximization (EM) algorithm, which can be used in cases with incomplete or missing data (Truxillo, 2005), was attempted in order to generate a covariance matrix. Later examination of the data showed that the coding of the variables in the American Community Survey (ACS) Public Use Microdata Sample (PUMS) files produced the missing values, not necessarily missing information. Once the coding was altered, the EM algorithm was no longer needed.

continuous and ordered categorical variables to ensure there was sufficient association amongst the variables. Consequently, a designated correlation matrix for this step in the process is not included. It was found, however, that 12 of the 16 variables correlated with a coefficient of at least 0.3,[53] which suggests reasonable suitability for factor analysis.

The following sections discuss the amended analysis on the random subsample, followed by a discussion of the findings. The chapter is then concluded.

## 7.3 2015 Exploratory Factor Analysis

From the 2015 ACS PUMS data, 16 variables are included. Prior to conducting the factor analysis, I must ensure the suitability and appropriateness of the data (Field, 2009). Two main issues are considered when determining the suitability of the data: the sample size and the strength of the correlation among the variables. The literature presents many rules of thumb for the minimum sample size needed to obtain a robust factor analysis (for instance, Yong and Pearce, 2013). The sample size for this data is 234,846, which is substantially large enough to conduct a factor analysis.

Three tests are recommended to determine sufficient correlation: a correlation matrix and two measures of sampling adequacy, the Kaiser-Meyer-Olkin (KMO) and the Bartlett test of sphericity. Section 7.2 confirmed that there is sufficient correlation amongst the variable to meet criteria. To proceed with the other tests, all nominal variables were transformed into binary variables. This step allows for factor analysis with the primary command in Stata, *factor*.[54] The transformed variables, presented in Table 7.1 are "means of transportation to work," "employment status," "industry of employment," "school enrolment" and "marital status." The binary variables created

---

[53] Sarstedt and Mooi (2014) note that all the elements of correlation matrices do not necessarily have to have high values. In the next section, I conduct other tests that further suggest suitability for factor analysis.
[54] This command generates a default correlation matrix in Stata. Binary variables take the form of 0/1, representing an interval level of measurement that can be used with the default Pearson's correlation in Stata.

represent each of their categories. To avoid multicollinearity, one category is eliminated from each of the transformed variables, leaving 49 variables (also called items) to be tested for both measures of sampling adequacy. Firstly, Bartlett's test of sphericity $X^2$ (1176) =0.00000221, $p<0.001$ is significant, indicating that correlations between the variables are sufficiently large for factor analysis. Secondly, the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy with all 49 items is 0.236. Kaiser and Rice (1974) note that any value below 0.50 is unacceptable. In order to enhance the suitability of the factor analysis, 29 items were removed as their individual KMOs were less than 0.3. Interestingly, these items were most of the binary variables representing 'means of transportation to work,' all of the 'industry of employment' indicators, and 'employment status: military.' The resulting KMO of 20 items was 0.736, which is an acceptable value. The Bartlett's test of sphericity was conducted with the 20 items, $X^2$ (190) =0.00000128, $p<0.001$. It is again significant, indicating that correlations between the variables are sufficiently large for factor analysis. Given the results of these tests, factor analysis was deemed to be suitable for the 20 of the original 49 items. Ferguson and Cox (1993) acknowledge that if satisfactory results are obtained with these two tests, the analyst can proceed with a factor analysis knowing that the matrix from the data is appropriate for factor analysis.

**Table 7.1: List of transformed variables for use in exploratory factor analysis**

| ACS Indicators for EFA | Dummy Categories |
|---|---|
| ***Means of transportation to work*** | drive (car, truck, van)<br>bus<br>streetcar<br>subway or elevated<br>railroad<br>ferryboat<br>taxicab<br>motorcycle<br>bicycle<br>walked<br>worked at home<br>other methods<br>not in the labour force (omitted) [55] |
| ***Employment status*** | employed civilian<br>employed military<br>not in the labour force/not seeking work<br>unemployed (omitted) |
| ***Industry of employment*** | agriculture<br>mining and other extraction<br>utilities<br>construction<br>manufacturing<br>wholesale<br>retail<br>transportation<br>information services<br>financial services<br>professional services<br>medicine<br>social services<br>entertainment<br>personal services, including unions<br>administrative<br>not in the labour force (omitted) |
| ***School enrolment*** | public education<br>private education<br>not enrolled in education (omitted) |
| ***Marital status*** | married<br>widowed<br>divorced<br>separated<br>never married (omitted) |

Source: 2015 ACS PUMS data dictionary

[55] Multicollinearity is also avoided with the elimination of newly formed binary variables that already capture necessary information. For instance, the 'industry of employment' and the 'employment status' variables both collect information about unemployment status. Information capturing an individual not being in the labour force and not seeking work only needs to be included once.

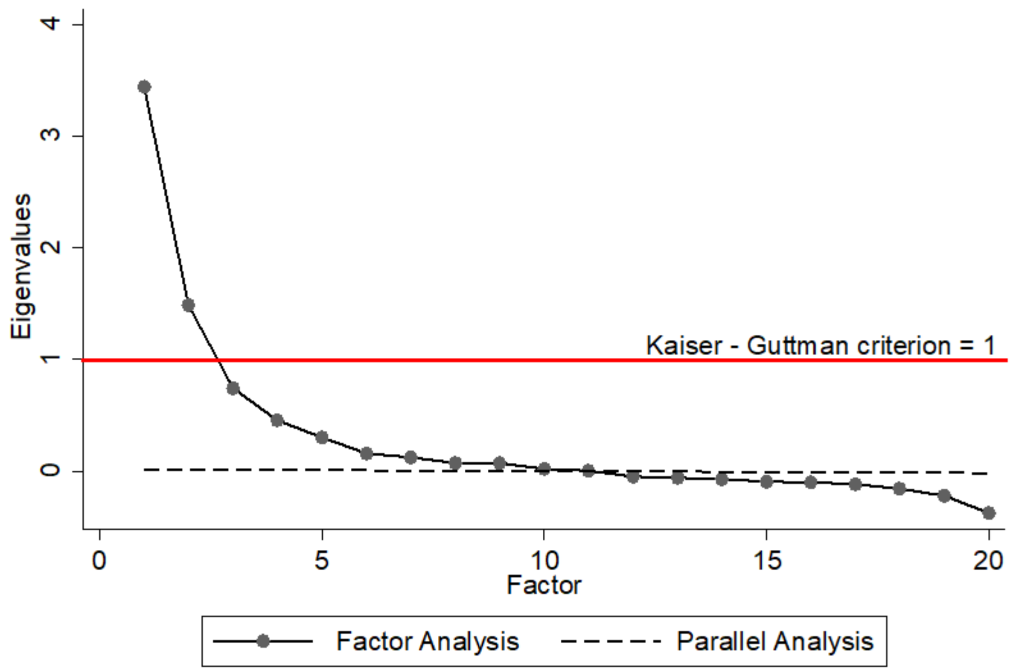
Section 6.2.2 addressed the decisions a researcher makes when conducting factor analysis, such as the estimation method, the proper rotation method, and the number of factors to retain (Costello and Osborne, 2005; de Winter and Dodou, 2012). As acknowledged in Chapter 6 (section 6.2.4), the estimation method selected is principal axis factoring (PAF), the rotation method is the oblique rotation, Promax, and the decision to retain factors is initially guided by parallel analysis.

I recognise that Yang and Xia (2015) caution against parallel analysis when the scale point is two, as found in binary variables. Because most of the final 20 items in this analysis are binary, I follow Matsunaga's (2010) advice to use parallel analysis as the primary method to determine the number of factors to retain in combination with qualitative scrutiny, particularly regarding the interpretability of factors and theoretical expectations regarding social exclusion.

An examination of the scree plot and parallel analysis shown in Figure 7.1 suggests that an 11-factor solution should be retained. Recall from the previous chapter (section 6.2.4) that parallel analysis suggests retaining any factors from the research data that are greater than the eigenvalues of the parallel analysis. Noting Costello and Osborne's (2005) recommendations for obtaining the cleanest (simple) factor structure, an 11-factor solution would not be ideal for several reasons. Firstly, an 11-factor solution yielded four factors with no items loading over 0.32, the minimum loading criteria recommended by Costello and Osborne (2005). Another four factors yielded one item each loading higher than 0.32 (‘marital status: widowed,’ ‘school enrolment: private education,’ ‘school enrolment: public education,’ and ‘marital status: separated,’ respectively). The recovery of weak factors (factors with one item) is more common in PAF than other extraction methods (de Winter and Dodou, 2012), so having these factors identified is not surprising. This does mean that in order to obtain the cleanest factor structure, it becomes necessary to adjust the factor retention method slightly.

Subsequently, the decision to retain factors was guided by inspecting the scree plot in Figure 7.1, while balancing the interpretability of factors. The visual inspection of the scree plot revealed that two factors had an eigenvalue greater than 1 and a bend after the fourth factor and another slight bend at the sixth factor. Methodologists Costello and Osborne (2005) specify that when the number of factors to retain is blurred, conducting a series of analyses is appropriate. Therefore, solutions for two, three, and four factors were examined. From the scree plot in Figure 7.1, it can be seen that after the fourth factor minimal variation is added to the construct. Accordingly, a five or six-factor solution is not considered further.

### Figure 7. 1: Scree plot of eigenvalues on 2015 EFA




Source: 2015 ACS PUMS
n=234,846

Of the 20 items included in the factor analysis, five items were eliminated at this stage because they did not contribute to a 2, 3, or 4-factor structure, failing to meet a minimum criterion of having a primary factor loading of 0.32 or above. Interestingly, these items are 'internet access,' 'English fluent household,' 'school enrolment: public education,' 'school enrolment: private education,' and 'marital status: separated.'[56] After deleting those five items, the 2, 3, and 4, factor structure was conducted again with the remaining 15 items. During this stage, one item, 'marital status: widowed,' was deleted, because it only loaded on a four-factor solution and was the only item loading on this factor. While there might be disadvantage and exclusion associated with widowhood,[57] this factor is not considered further as factors with one item are considered weak and ill-defined (Costello and Osborne, 2005).

### Table 7. 2: Factor Loadings of 2015 EFA

| Factor | Eigenvalue | Difference | Proportion | Cumulative |
|---|---|---|---|---|
| Factor 1 | 3.27448 | 1.95110 | 0.7242 | 0.7242 |
| Factor 2 | 1.32337 | 0.75589 | 0.2927 | 1.0169 |
| Factor 3 | 0.56748 | 0.41353 | 0.1255 | 1.1424 |
| Factor 4 | 0.15396 | 0.05062 | 0.0340 | 1.1764 |
| Factor 5 | 0.10334 | 0.08581 | 0.0229 | 1.1993 |
| Factor 6 | 0.01753 | 0.05187 | 0.0039 | 1.2031 |
| Factor 7 | -0.03434 | 0.02140 | -0.0076 | 1.1955 |
| Factor 8 | -0.05574 | 0.00538 | -0.0123 | 1.1832 |
| Factor 9 | -0.06112 | 0.00947 | -0.0135 | 1.1697 |
| Factor 10 | -0.07059 | 0.03273 | -0.0156 | 1.1541 |
| Factor 11 | -0.10332 | 0.05274 | -0.0229 | 1.1312 |
| Factor 12 | -0.15606 | 0.03924 | -0.0345 | 1.0967 |
| Factor 13 | -0.19529 | 0.04678 | -0.0432 | 1.0535 |
| Factor 14 | -0.24207 | . | -0.0535 | 1 |

Source: American Community Survey, 2015 PUMS
N: 234,846

[56] Some of these items made up of the weak factors in the 11 factor solution discussed previously.
[57] Potential explanations for these items not loading and not subsequently considered in the analysis are explored later in this chapter (section 7.3.1).

For the final stage, another PAF factor analysis was conducted with the remaining 14 items using a Promax rotation. Three factors explain over 100% of the variance,[58] shown in Table 7.2. The three-factor solution was preferred over a two factor solution for two reasons: 1) the factor loadings load decisively allowing for a definitive and substantive interpretation as recommended by Matsunaga (2010)[59] and 2) it has been suggested that the under extraction of factors is more problematic than the over-extraction, as it leads to a loss of relevant information (Ledesma and Valero-Mora, 2007; Wood *et al*., 1996). Beavers *et al.* (2013) recognise a consistent caution against a one or two-factor solution as it may not provide an accurate representation of the structure under investigation. All of the remaining 14 items in factor analysis had primary loadings over 0.32, none of which cross-loaded. The factor loading matrix for this final solution (after rotation) is presented in Table 7.3.

[58] An oblique rotation method was used for this analysis, which allows the factors to correlate. Allowing the factors to correlate mean that each of the factors share variance of the underlying construct. As a result, this number does not indicate the true amount of variance of the underlying construct, social exclusion, explained by the retained factors. This number is over estimated.

[59] In this analysis, the item 'marital status: married' loads on the second factor in a two factor solution. However, in a three factor solution, the item loads strongly in the third factor, along with two other similar items, 'marital status: divorced' and 'unmarried partners,' allowing for substantive interpretation.

**Table 7.3: Obliquely rotated factor loadings for EFA on 2015 ACS PUMS data**[60]

| Factor | Labour Force Participation | Economic Security | Marriage: a Social Resource[61] | Uniqueness |
|---|---|---|---|---|
| Disability | -0.34 | | | 0.86 |
| Health Insurance | | 0.35 | | 0.89 |
| Overcrowded housing | | -0.39 | | 0.87 |
| Relative income poverty | | -0.57 | | 0.60 |
| Travel time to work | 0.57 | | | 0.66 |
| Not in Labour Force | -0.94 | | | 0.12 |
| Employed civilian | 0.95 | | | 0.07 |
| Means of transportation to work: drive | 0.82 | | | 0.30 |
| Educational attainment | | 0.50 | | 0.74 |
| Marital status: married | | | 0.58 | 0.60 |
| Marital status: divorced | | | -0.53 | 0.73 |
| Unmarried partners | | | -0.35 | 0.85 |
| Total individual income | | 0.45 | | 0.75 |
| Food stamps | | -0.43 | | 0.80 |
| Variance | 3.14 | 1.74 | 0.95 | |
| Proportion of variance | 0.70 | 0.38 | 0.21 | |

Loadings >0.32 are shown.[62]
Source 2015 ACS PUMS
n=234,846

Some of the commonalities[63] shown in Table 7.3 are below the often cited threshold in the social sciences of 0.4 (Costello and Osborne, 2005). This is likely due to the large sample size used for the factor analysis. It has been acknowledged that it is possible to recover population factors and obtain good factor congruence when commonalities are low if the sample size is extremely large (Hogarty *et al.*, 2005; MacCallum *et al.*, 2001, 1999). Given that most discussion in the relevant literature focuses on small sample sizes (for instance, Jung, 2013) and sample sizes of 200 are

[60] I also ran an EFA for the 2008 ACS PUMS. A table depicting the obliquely rotated factor loadings for that EFA is available in Appendix B. The results of the EFA for 2008 are consistent with the loadings for 2015, suggesting that using the available indicators for this group of the population (individuals aged 18 and up), the factors of disadvantage were consistent in the United States before the Financial Crisis and after.

[61] Because the direction of the factor loadings for the 'marriage as a social resource' factor was opposite from the other two, the loadings were multiplied by -1, so that an increase in the factor is associated with more advantage. This enables a consistent discussion of disadvantage and advantage across the factors.

[62] In section 6.2.4, I acknowledged that I retain factors with loadings of at least 0.32 as recommended by Costello and Osborne (2005).

[63] The communalities are equalled to 1 – uniqueness. The uniqueness for each factor is presented in Table 7.3.

considered large (for instance, Jung and Lee, 2011), we can be sure that the sample size used here can be described as 'extremely large'. Therefore, despite having low communalities, the factor structure is accepted as presented in Table 7.3.

Five items load on factor 1. Table 7.3 clearly shows that these items relate to work and employment. The items loading on to this factor are 'disability status,' 'travel time to work,' 'employment status: employed civilian,' and 'driving as a means of transportation to work'. This factor is labelled "*Labour Force Participation*." The internal consistency for this factor was tested using Cronbach's alpha, the value of which for this factor is 0.85. As acknowledged in the previous chapter, there is no set cut-off value for Cronbach's alpha (Schmitt, 1996), but George and Mallery (2003) suggest that anything above 0.5 is an acceptable value. With 5 items loading on this test, a value of 0.85 represents a high variance consistency. Essentially, this indicates that the items in this factor have shared covariance and are likely measuring the same underlying construct. As recommended by Schmitt (1996), Bartlett's test of sphericity is also included to ensure consistency of this factor. The null hypothesis of the Bartlett test is that the variables are not inter-correlated. With an $X^2$ (10) =0.000001, p<0.001, we can be 99% confident that the items within this factor are inter-correlated.

Six items load on factor 2: 'Health insurance,' 'living in overcrowded housing,' 'total individual income,' 'income poverty status,'[64] 'food stamps recipient,' and one's 'educational attainment.' These items relate to "*Economic Security*" and this factor is labelled as such. The Cronbach alpha for this factor is 0.6061. The value of this Cronbach alpha specifies that there is an above-average high variance consistency. With an $X^2$ (15) =0.00000131, p<0.001, there is a 99% confidence that these variables are inter-correlated.

Three items load onto factor 3. Table 7.3 shows that these items are related to social resources and contact. The items that load on this factor are 'marital status: married,' 'marital status: divorced,' and 'unmarried partners.' This factor is labelled '*Marriage as a Social Resource.*' The

[64] As defined by this thesis: 250% below the poverty line.

Cronbach alpha for this factor is 0.5457. Tavakol and Dennick (2011) acknowledge that alpha is a function of the covariance amongst the items and the number of items in a factor. Therefore, when the number of items in a test is small, the value of Cronbach's alpha is also likely to be small. Tavakol and Dennick (2011) suggest increasing the number of items in the test to increase alpha. This recommendation was not taken for several reasons. Firstly, there is no consistent recommendation for the cut-off for alpha, because it is a function of a number of things. Secondly, Schmitt (1996) notes that low alphas do not correspond with low reliability. Finally, qualitative examination and substantive interpretation are relied on for this factor analysis. The three items loading on this factor meet that criteria. Importantly, there are no additional and relevant items available within the PUMS data set that could substantially load onto this factor. This value for Cronbach's alpha does indicate that over 50% of the variance in this factor is consistent and these items are related. With an $X^2$ (3) =51807.095, $p<0.001$, there is a 99% confidence that these variables within this factor are inter-correlated. Consequently, this factor structure is utilised as a domain of social exclusion, maintaining that a factor with 3 items is sufficient (Costello and Osborne, 2005).

## 7.3.1 Some items did not load. Why not?

In the various stages of the exploratory factor analysis, there were six items that either did not load high enough to be included in the final structure or loaded in one factor. These were 'internet access,' 'English fluent household,' 'school enrolment: public education,' 'school enrolment: private education,' and 'marital status: separated', and 'marital status: widowed.' In this section, I explore why some of these items may not have loaded strong enough on any of the factors.

Firstly, the variable 'internet access' did not load. Studies exist regarding the relationship between internet usage and well-being, but they produce mixed results (Boniwell *et al.*, 2015). In addition, there are differences between internet use and internet access, the latter of which is analysed here. The data from the 2015 ACS PUMS focuses on the digital divide, which refers to the binary distinction between having and not having access to the internet (DiMaggio and Hargittai, 2001). Many in the literature have suggested that it is necessary to move away from this focus to digital inequality in terms of quality and quantity (Boniwell *et al.*, 2015; DiMaggio and Hargittai, 2001; van Deursen and van Dijk, 2014), but also with a focus on the influence that internet usage has on social structure (Hampton, 2010). The results here support the relevant literature, such that continued focus on the digital divide is misguided.

'English fluent household' also failed to load on any factor. There are social and economic reasons why language can contribute to disadvantage, particularly for immigrants. On the social side, speaking English poorly may result in discrimination and contribute to social isolation (Treas and Mazumdar, 2002). The economic reason is where most of the literature about English speaking ability resides, particularly its relationship to the labour market (Bloom and Grenier, 2009; Chiswick and Miller, 2002, 1995; McManus, 1990). The literature does offer a solution about why this item may not have loaded. It has been suggested that immigrants tend to stay within labour market enclaves. Focusing entirely on Hispanic men in the United States, McManus (1990) notes that these enclaves provide better jobs for those with limited English speaking ability. Immigrants tend to move to linguistic concentrated areas (Chiswick and Miller, 2002), making them less disadvantaged in their own communities.

Both items relating to school enrolment, public and private education, did not load on any factor. This variable from the ACS PUMS was included to capture future advantage believed to be acquired via education. Recall that this analysis includes individuals aged 18 and up. It is likely that most individuals who are enrolled in school (though the data does not specify if it is enrolment in grade school, secondary school, or university) are not in this age group. In addition, the role of education in examining multidimensional disadvantage in the United States is likely captured in the item 'educational attainment,' which loaded in each iteration of the factor analysis.

'Marital status: separated' also did not load on any of the factors. In much of the literature surveyed to determine possible reasons for this item not loading, divorce is often discussed in tandem with separation, even combining the two as one (Levinger, 1976; Martin, 2006; Thornton, 1985).

Separation, like divorce, represents a dissolution of a union. Consequently, it is quite likely that an individual who separates from their marriage partner experiences similar social and economic effects, including a declined standard of living (McManus and DiPrete, 2001), as does an individual who is divorced. The item, 'marital status: divorce' did consistently load in the various iterations of the factor analysis. This item is likely carrying effects for both variables, making 'marital status: separated' redundant.

The last item deleted from the analysis was 'marital status: widowed.' What separates this item from the others above is that it actually loads in a four-factor solution. Because factors with just one item are considered ill-defined (Costello and Osborne, 2005), it was not included in the analysis. There are exclusion and disadvantage associated with widowhood with which the other items included in the factor analysis would not strongly group. For instance, Osberg and Sharpe (2014) acknowledge that there is a hazard associated with widowhood as the underlying event, death, happens to someone else (the spouse or partner), with whom the widow had her fortunes linked by marriage. Indeed, there is an implicit context associated with this statement and for an increasing amount of women, marriage does not guarantee stability. However, many of the women under analysis here are in retirement age or are facing retirement age[65] and spent their adulthood in that implicit context in which a women's security was associated with and depended upon her husband's (Angel *et al.*, 2007). Due to the time-limited nature of this project, this aspect of disadvantage in the United States is not further explored. However, this is likely to be a key issue for the Baby Boomer generation as they grow older.

[65] In the 2015 ACS PUMS data, 13% of the sample are retirement aged women (65 and up). 10% of the sample are near retirement aged, between the ages of 55 and 64.

## 7.4 Substantial interpretation of the three factors of (dis)advantage in the United States

The analysis above highlights three factors of (dis)advantage. Due to the direction in which the individual items load on each factor, these factors are labelled advantages rather than disadvantages. The dimensions of advantage in the United States are 'labour force participation,' 'economic security,' and 'marriage as a social resource.' The following sub-sections explore the theoretical understandings of each of these factors.

### 7.4.1 "Labour force participation"

Participation in the labour force has long been discussed in tandem with poverty in the United States (McKeever and Wolfinger, 2011; Mooney, 1967; Wilson, 1987). Being in employment is often lauded as the best route out of poverty, toward integration, and to escape disadvantage. This is undoubtedly true in the United Kingdom where increasing the labour force participation rate has been used as a key policy objective to decrease social exclusion (Levitas *et al.*, 2007). The attention to the labour force participation rate is based on traditional notions of poverty that focus narrowly on income and based on the premise that most individuals derive their income from work. Nonetheless, the literature has acknowledged that the focus on the unemployed poor and the relevant disadvantages is not a complete discussion related to the institution of the labour market (Brady *et al*., 2013; Leach *et al*., 2010).

Based on the findings uncovered in the factor analysis, labour force participation is indeed a key factor of (dis)advantage in the United States. However, the analysis suggests that labour force participation as a factor of (dis)advantage has two components: 1) actual participation and 2) the ability to participate.

In the first understanding is the traditional and basic understanding of labour force participation. That is, whether or not an individual is actually in the labour force. Two of the five items loading on this factor relate directly to that: 'employment status: employed civilian'; and 'employment status: not seeking work.' The latter item highlights the voluntary nature of exclusion in that a person can choose to self-exclude them self from this piece of society.

Given the changing structure of the labour market and of work, the analysis highlights that another aspect deserves simultaneous attention: the ability to participate in the labour force. This is seen with the additional items that load onto this factor: 1) 'disability status,'2) 'travel time to work,' and 3) 'means of transportation to work: drive.' Individuals of working age with a disability are recognised to be more likely to live in income poverty than those without (Stapleton *et al.*, 2006; Yeo and Moore, 2003).There are often conditions of the labour market that exclude individuals with disabilities[66] from actually participating in the labour market. For instance, Stapleton *et al.* (2006) recognise that individuals with disabilities are often not able to work as they are trapped in policies that provide benefits on the condition of not being able to work. These benefits tend to be lower than what one would receive if they were unemployed (Burkhauser *et al.*, 2016). The disadvantages faced by individuals with disabilities have a long history, as highlighted by many legal rulings to address discrimination and exclusion (Krahn *et al.*, 2015). However, the finding here and elsewhere (Krahn *et al.*, 2015) suggest a need to explore this and its connections to exclusion. This thesis finds a particular connection between labour force participation (dis)advantage and disability status.

Another item in the 'labour force participation' factor that centres this notion of ability to participate is 'travel time to work,' which interestingly loaded higher on this factor than disability status. The time in which it takes an individual to travel to work centres on spatial relationships to the labour market and the subsequent consequences of these relationships for social exclusion. The most recognisable concept discussed within the literature that relates is the spatial mismatch hypothesis, which posits that urban employment levels are affected by shifting job and residence locations

[66] It is important to restate that disabilities in this thesis refers to difficulty in vision, hearing, cognition, and/or ambulation.

(Sanchez, 1999). What underlies the spatial mismatch hypothesis is that there is a scattering of jobs away from central cities within the United States (Kain, 1968; Sanchez, 1999). Holzer (1991) acknowledges that the distance between a central city resident and likely work locations have been increasing over time.

Some have argued that commute time to work may be a misleading measure of spatial mismatch (Immergluck, 1998; Morrison, 2005). There are two things to note in response to this. Firstly, the relationship between the 'travel time to work' item and the labour force participation factor was positive. This suggests that it is important in discussing a person's ability to work as the increase in travel time is associated with an increase in the ability to participate in the American labour force. Secondly, it has been noted that as the typical distance needed to travel from home to work increases, those with low skills and low levels of personal mobility (such as the use of personal car) are not able to meet the travel requirements of dispersed employment increases (Sanchez, 1999), which tends to impact and disadvantage minorities more so than White Americans (Easley, 2018; Kain, 1968). The analysis uncovered here lends weight to the hypothesis, particularly as 'means of transportation to work: drive'[67] also loaded on to this factor, higher than any of the items representing ability to participate in the labour force. Morrison (2005) acknowledges that access to a car can widen employment horizons. Therefore, a person who is capable of longer commute times is much more likely to find employment. This then relates to other factors regarding an individual's circumstance or constraints that prohibit them from making long commutes. For instance, if an individual relies solely on public transportation, they are not able to explore job opportunities beyond their regional and local area, as bus routes are geographically limited. In addition, mothers - particularly single mothers - who work and provide primary care for children may not be able to make

[67] Of the available options from the PUMS data for 'means of transportation to work', driving is the only one that met initial criteria to be included in the analysis. As such public transportation is not included. This analysis is not suggesting that access to public transportation is not important. Rather, it suggests that for the United States as whole, driving is a key component of advantage in labour force participation. Had this analysis been conducted for a central city, the factors may differ slightly and support others, such as Sanchez (1999), who argue that public transport is important to improve urban employment.

a longer commute to work, and subsequently may remain in underemployment or take on a lower-paying job in closer proximity.

Considering solely the rate of employment in society is not a reliable indicator of the reasons that can combine to exclude an individual from full participation in the labour market. It has become evident that we have to also understand participation behaviour and recognise the constraints to employment. Taking two components of labour force participation and analysing them concurrently, this factor offers a complete picture of (dis)advantage in labour force participation in the United States.

### 7.4.2 "Economic security"

The use and perception of the term 'economic security' have seen a rise among the world's wealthiest nations since the Financial Crisis that began in December of 2007. Much of the literature on the topic focuses on the lack of security (for instance, Bossert and D'Ambrosio, 2013; Hacker *et al.*, 2014; Western *et al.*, 2012; Wroe, 2016). The analysis I have undertaken in this thesis, however, suggests an interpretation as an advantage. The relevant literature is still useful in providing meaning and interpretation of this factor.

Economic (in)security is a multifaceted issue and a formal, comprehensive definition is difficult to ascertain. Some have described it as objective (Hacker, 2011) and some subjective (Dominitz and Manski, 1996), while others conceive it as a combination of the two: perception and actual risk (Jacobs, 2007). The United Nations (2009) recognises the difficulty in defining and measuring economic insecurity but offers a definition, suggesting that *"economic insecurity rises from the exposure of individuals, communities, and countries to adverse events and from their inability to cope with and recover from the costly consequences of those events"* (quoted in Bossert and D'Ambrosio, 2013, p. 1018). This two-fold understanding of economic (in)security, the exposure to events and the inability to recover from those events, adequately summarises the items loading on this factor.

One on hand, three of the six items loading on this factor: 1) 'living in overcrowded housing,' 2) 'income poverty,' and 3) being 'food stamps recipient,' reflect the acknowledgement that economic insecurity arises when individuals are exposed to adverse events. Table 7.3 shows that each of these items has a negative relationship with this factor. These adverse circumstances put individuals at risk of social exclusion in the form of economic insecurity. The analysis suggests that should an individual live in overcrowded housing, it could be quite possible she does not have enough economic resources to provide adequate housing. Living in overcrowding housing is linked to other adverse events, such as food insecurity (Liu *et al.*, 2014), which is captured in the item 'food stamps recipient,' which also loads negatively on this factor. Recipients of food stamps do not have a lot of economic resources and are entitled to supplemental benefits from their respective state government to provide food for themselves. Food insecurity is linked to detrimental impacts on health for adults and children (Cook *et al.*, 2013). Being in income poverty ties this together because if an individual does not have adequate income, they are less likely to have secure housing and adequate access to food.

On the other hand, the United Nations (2009) acknowledge that economic insecurity arises when individuals are unable to deal with and recover from these adverse events. The remaining three factors further encompass this understanding of economic insecurity. Firstly, having health insurance is recognised to be a key component of economic security (Dhongde and Haveman, 2016; Haley and Rejda, 2001; Western *et al.*, 2012). Indeed, there are noted economic consequences of ill health or an accident that requires great medical attention. Hacker (2007) understands that high rates of underinsurance and non-insurance works with the rise in costs to substantially increase income risks for the middle class. People with high health care costs, particularly for the uninsured and those with chronic conditions, often borrow on credit to pay (Western *et al.*, 2012). As a result, reducing the percentage of the uninsured in the United States is often suggested as a method to reducing major causes of economic insecurity (Haley and Rejda, 2001). Educational attainment and total individual income both have positive relationships within this factor. We would expect that increases in educational attainment are associated with better paying job opportunities, increasing income,[68] thereby providing a greater degree of economic security. This two-fold understanding of economic (in)security fits well with the results of the factor analysis.

### 7.4.3 "Marriage as a social resource"

There is an understanding that marriage tends to provide benefits that are not enjoyed by the unmarried (Wells and Zinn, 2004). For instance, marriage is often suggested as a key to economic well-being and avoidance of poverty (e.g. Amato and Maynard, 2007). Family structure is so entrenched in the American understanding of poverty that the official measure is linked to it (Thiede *et al.*, 2017). An individual is considered poor if their family's pre-tax cash income falls below a threshold that is defined according to family size, age of the householder, and the number of children in the household under the age of 18 (Orshansky, 1965b). In addition, the 1996 welfare legislation, Temporary Assistance for Needy Families (TANF), "positioned the ethic of 'strong families' and the importance of marriage as fundamental to the prevention of poverty and 'welfare dependency'" (Geva, 2011, p. 25).

Perhaps due to these ties between family structure and poverty, there exists a large body of research that examines the connections between family structure and economic prosperity (e.g., Baker, 2015; Prokos and Keene, 2010; Seccombe, 2000; White and Rogers, 2000; Zagorsky, 2005). Indeed, from an economic perspective, there are many gains from marriage, such as the division of labour which increases a couple's productivity, economies of scale (such as cooking one meal together instead of two separate meals), and the pooling of resources and risks (Lehrer, 2008). However, viewing marriage through this limited lens often ignores various perspectives.

[68] Hacker (2007) has noted that the incomes of Americans have been on a 'rollercoaster,' as they rise and fall much more quickly now than they did a generation ago across the educational spectrum. This in itself is a form of insecurity.

Viewed another way, the connections between marriage and disadvantage can be linked to notions of social capital.[69] Lin (1999) argues that social capital is captured from rooted resources in social networks. Along these lines, indeed, the family is an essential source of social capital (Furstenberg and Kaplan, 2004). At the individual level, Johnson *et al*. (2011) recognise that social capital can produce two positive externalities, social support and social leverages. For instance, family structure can provide a foundation of health-related attitudes and behaviours (Grzywacz and Fuqua, 2000). People in dual-earning marriages can modify intra-household choices, provide and gain access to health care, and expand social networks (Grzywacz and Fuqua, 2000).

The social support system obtained from unions can be seen further via an examination of the results of the factor analysis in Table 7.3. Individuals who are in an unmarried-partner household have a negative relationship with this factor. However, a divorced individual is less advantaged than an unmarried partner. Bumpass *et al.* (1991) recognise that unmarried partners tend to be less stable than married partners, as their union tends to be a new arrangement, indicating that social leverages coming with marriage have yet to be attained. The only relationship in this factor that is positive is for individuals who are married. This suggests that marriage does indeed provide an advantage,[70] promoting the idea that we can frame marriage as a social resource and not solely an economic one. It could be that the social leverages in a married household have had time to develop in order to engage those social resources. The ACS PUMS does not include information about the number of

[69] Levitas *et al.* (2007) prefer the term social resources over social capital due to the latter's focus on the community over the individual. Social capital can be focused at the individual level as seen in the work of Coleman (1988) who specifies capital outside and within the family. In addition, both the definitions for both social resources and capital suggest these can be examined in tandem. Donenfeld (1940) define social resources "as the organised associations, institutions, and attitudes, both public and private, which society has developed to satisfy its physical, psychologic, economic and social needs" (p.562). Putnam (1995) defines social capital as features of social organisation, such as networks, norms, and social trust that facilitate coordination and cooperation for mutual benefit (p. 67). All these definitions highlight social networks and actors who operate within those networks. In essence, individuals are able to attain advantages via a variety of social relationships operating within various institutional domains that make up a society (Kabeer, 1999).

[70] I recognise that the relationship between marriage and advantage is not causal. There are definitely mechanisms at play that make a union successful, which could include age at marriage (Lehrer, 2008) and/or religious heterogamy (Heaton, 2002). An exploration of these factors is not within the scope of this thesis, but they are areas suitable for future research.

years an individual is married or has lived in an unmarried-partner household. Consequently, those connections cannot be explored further here. However, this thesis examines marriage and its connection to disadvantage as a lack of family-based social resources. The significance of exploring this factor in such a way is that we are mostly concerned with family structure according to its functions, and as Coleman (1988) acknowledges, how individuals use these functions to achieve their interests.

In my interpretation of this factor, I do recognise and acknowledge that the structure of the family in the United States has gone through a significant transformation in the past fifty years (Prokos and Keene, 2010) and relates to what Cherlin (2004) dubs the deinstitutionalisation of marriage. This raises important questions about the relationship between family structure and the continued reliance placed on marriage in the United States in order to decrease income poverty rates. The results of the factor analysis presented in Table 7.3 highlight that the item 'unmarried partners' is an integral part of this factor.[71] Others have recognised that cohabitation is likely contributing to the decline in marriage rates in the United States (Bumpass *et al.*, 1991; Cherlin, 2004). In the past, cohabitation was often more experienced by the least educated of the American population and the income poor (Bumpass *et al.*, 1991). More recently, cohabitation is more readily accepted as an alternative to marriage (Kiernan, 2002; Smock and Gupta, 2002). But differences in the processes of cohabitation and associated family outcomes diverge by social class (Sassler and Miller, 2011). For instance, Lichter *et al*. (2006) acknowledge that cohabitation is substantially more likely to lead to marriage for nonpoor women, whereas economically disadvantaged women tend to cohabitate in succession (Sassler and Miller, 2011). Economically advantaged individuals, for instance, who are likely to not be socially excluded, tend to progress into cohabitation after having a stable relationship for over a year (Sassler and Miller, 2011). Sassler and Miller (2001) also acknowledge that working class individuals have reported housing and other practical financial reasons to cohabitate, whereas

[71] The 'never married' indicator was not included in the factor analysis as it served as the reference category for marital status.

middle class, educated individuals tend to use cohabitation as a step to marriage and a test of the suitability of the relationship. This highlights that for those least economically disadvantaged and least likely to be socially excluded tend to use cohabitation for different reasons. It also suggests that the factor itself is likely not to be experienced or utilised similarly amongst various groups of the population.

Exploring this factor as a social resource, there are some caveats worth addressing as some individuals will not fit neatly into the interpretation highlighted above. Firstly, in discussing 'marriage as a social resource,' it is important to consider the benefits of marriage whilst in poor relationships. Research suggests that married couples enjoy better health outcomes and social support (Cutrona, 1996). However, the same social ties found in marriage that are to be sources of support can also cause stress and conflict (Horwitz *et al*., 1998), thereby decreasing the benefit of marriage as a social resource. While Gove *et al*. (1983) suggest that marriage itself tends to be more important for the well-being of men and the quality of the marriage is more important for women, others have suggested that poor marital quality has negative effects for both in a union (Johnson *et al*., 2000). Poor marital quality then suggests that individuals within the union are unable to utilise the positive externalities associated with social relationships, particularly social support.

Secondly, I recognise the need to explore self-exclusion from marriage. There are a variety of reasons why people choose not to marry, making this particular social resource less desirable (Donenfeld, 1940). This may relate to the choice to either remain single, delay marriage, or remain in prolonged cohabitation. For instance, research has shown that in areas in which there is relatively high male wage inequality, women tend to search longer for husbands (Gould and Paserman, 2003). Additionally, Hatch (2017) suggests that there is a subset of cohabiters that actively resist marriage and living together is not a test for marriage. They do acknowledge that this subset of unmarried partners resist marriage because marriage may conflict with their ideals in relation to equality and civil liberties (Hatch, 2017). Individuals who are not likely to be socially excluded - the economically advantaged and educated, for instance - choose not to engage with this particular dimension of 'inclusion' and 'advantage' for various personal reasons.

The findings for this factor and its subsequent interpretation suggest that various considerations should be made in discussing its role to disadvantage. For instance, marriage as a social resource may not be a desirable resource to attain. In addition, the selection into marriage also varies by class. It is my hope that by exploring these different avenues, this factor does not revert to suggesting that pro-marriage campaigns aid in increasing marriage among the poor and elevates couples out of income poverty (Lichter, 2001; Rector, 2010).

## 7.5 The influence of age on the factor results

A decision was made in Chapter 5 (section 5.3.1) to include all individuals aged 18 and up in this analysis. There is merit in conducting an analysis on this broad range as I can determine what the factors of disadvantage might be for the entire adult population in the United States. With that, I recognise that the factors of (dis)advantage uncovered in this analysis may be experienced differently by different age groups and the meaning of these factors may vary by age, particularly for those aged 65 and up. Indeed, this is a key research question for this thesis and the analysis in Chapter 8 will highlight that individuals aged 65 and up are less advantaged in 'labour force participation', on average, compared to the other age groups.

The indicators, however, within this factor may highlight some of these differences. For instance, two of the items loading on 'labour force participation: 'employment status: not looking for work' and 'disability status' may be more likely to be experienced by older individuals who are retired. While non-participation in paid work has been recognised as a component of social exclusion as acknowledged in Chapter 2 (particularly, in discussions of the discourses of social exclusion), individuals who are over 65 may not necessarily be socially excluded due to retirement. An individual who is over 65 may have disabilities that are not related to their non-participation in paid work; their disability itself may be age related. I recognise that these indicators may not be completely generalizable to 'labour force participation' for all age groups.

With this acknowledged, I keep individuals aged over 65 in the analysis for one primary reason. The analysis in this thesis argues that labour force participation is not the sole factor of note, even if individuals aged over 65 may have a different relationship with this factor than the working age population. Considering the effect this decision might have on the results of the retained dimensions of (dis)advantage, I conducted a factor analysis on the working age population (ages 18-64) and the factors remained approximately the same. The results provide some confidence that age is not heavily skewing the uncovered dimensions of disadvantage in the United States. This also provides further evidence that labour force participation is not the only dimension of social exclusion in the United States. Like Dhongde and Haveman (2016), I am generally interested in exploring these differences and deriving these factors based on the entire population.

## 7.6 Using the results of the factor analysis in subsequent analysis

In Section 6.2.2, I acknowledged that an additional purpose of deriving the factors of (dis)advantage is for their ability to be used in additional analysis. Factor scores, as acknowledged in Chapter 6, can be derived from each factor separately to develop continuous dependent variables. The factors extracted from the EFA will provide continuous measures of different aspects or dimensions of social exclusion. The dependent variables are utilised to analyse the association between each factor of (dis)advantage and individual sociodemographic characteristics. The descriptive statistics for the factor scores,[72] generated via Bartlett factor score regression method (addressed in section 6.2.2.) are found in Table 7.4.

### Table 7.4: Summary of EFA factors

| Factor | Name | Count | Mean | Std. Dev. | Min | Max |
|---|---|---|---|---|---|---|
| 1 | Labour Force Participation | 234,846 | 0 | 1.03 | -1.39 | 1.23 |
| 2 | Economic Security | 234,846 | 0 | 1.28 | -5.15 | 8.45 |
| 3 | Marriage as a Social Resource | 234,846 | 0 | 1.42 | -5.03 | 2.16 |

Source: 2015 ACS PUMS

Table 7.5 presents the correlation matrix for each of the factors. The highest correlation exists between the 'labour force participation' and 'economic security' factors. The shared variance for these two factors is 4.8%. The 'marriage as a social resource' factor has a low correlation with both of the other factors. The shared variance with 'labour force participation' is less 0.09% with 'labour force participation' and 1% with 'economic security.'

### Table 7.5: Correlation matrix of the three factors of (dis)advantage

| | Labour Force Participation | Economic Security | Marriage as a social resource |
|---|---|---|---|
| Labour Force Participation | 1 | | |
| Economic Security | 0.22 | 1 | |
| Marriage as a Social Resource | 0.03 | 0.1 | 1 |

Source: 2015 ACS PUMS

[72] Typically, the factor scores have a mean of zero and a standard deviation of 1. The standard deviation for the 'marriage as a social resource' dimension deviates the most from this. It seems that this value is influenced by outliers among the states. This is addressed in Chapter 9 and depicted in Figure 9.3.

The results in Table 7.5 indicate that each of the three factors of (dis)advantage in the United States are correlated, yet distinct concepts that can be analysed separately. Each will be examined in the following chapter, analysing their respective associations with individual sociodemographic characteristics.

## 7.7 Conclusion

The purpose of this chapter was to answer the first research question ***(What is/are the factor(s) of multidimensional disadvantage in the United States?)*** By applying American indicators to the Bristol Social Exclusion Matrix, an exploratory factor analysis uncovered three factors of (dis)advantage: labour force participation, economic security, and marriage as a social resource. The analysis lends weight to the argument that (dis)advantage is not a unidimensional phenomenon. Income has shown to be a small piece of a complex social issue. The implications for these findings are further discussed in Chapter 10.

# Chapter 8

# The relationships between individual characteristics and multidimensional disadvantage

## 8.1 Introduction

In Chapter 2, I recognised that identity could form a basis for disadvantage (Jackson, 1998; Kabeer, 2000; Purdie-Vaughns and Eibach, 2008). An individual can experience disadvantage in society just by virtue of who they are. Subsequently, the purpose of this chapter is to examine the association between individual sociodemographic characteristics and multidimensional disadvantage in the United States. The sociodemographic characteristics of interest are age, race, gender, citizenship status, and the intersection between gender and race, as discussed in section 2.4. In the methods chapter (Chapter 6), an Ordinary Least Squares (OLS) multivariate regression analysis was identified as the optimal method to explore these relationships. It is chosen because it allows more than one independent variable to be present in the models, allowing for the concurrent examination of the relationships between the independent and dependent variables. This will provide insights into whether different groups experience disadvantage differently. By applying this statistical technique, this chapter will answer the following research question:

*To what extent are sociodemographic characteristics associated with multidimensional disadvantage?*

## 8.2 Replicate weighting in large scale data: addressing sampling error

I recognise that because the data used for this analysis (the Public Use Microdata Sample file) are derived from the full version of the 2015 American Community Survey, it may have some sampling error. In order to correct some issues related to this, the Census Bureau (2014) suggests the use of replicate weights, which have been available for the ACS PUMS files since 2005. Replicate weights allow PUMS users to generate empirically derived standard errors (Census Bureau, 2014), which can then be used to construct confidence intervals around the individual sociodemographic characteristics under investigation in this thesis. The replicative weights also allow the sample to be representative of the American population. Subsequently, all OLS analyses undertaken utilise the sampling weights (80 replicate weights) to account for the complex survey design. This can be done using the *svy* collection of commands in Stata 15. This produces a population size of 23,935,012. There are no missing data for the variables of interest, which were addressed in Chapter Five (section 5.3). Accordingly, the analyses continued with complete cases. Table 8.1 presents the weighted percentages (along with confidence intervals) for the non-intersectional sociodemographic variables in the 2015 ACS PUMS sample. For each category, the most represented are Female, White, aged 65 and up, and citizenship status: born in the U.S. This paints a compelling portrait of the typical American in 2015 and provides an interesting starting point at exploring who may be individually disadvantaged in the United States.

**Table 8.1: Weighted percentages of the sociodemographic variables**

| Sociodemographic Characteristics | Percentage in sample (95% Confidence Interval) |
|---|---|
| ***Age*** | |
| 65 and up | 19.19% (19.04% - 19.34%) |
| 55-64 | 16.81% (16.64% - 16.97%) |
| 45-54 | 17.59% (17.38% - 17.81%) |
| 35-44 | 16.89% (16.71% - 17.08%) |
| 25-34 | 17.73% (17.53% - 17.93%) |
| 18-24 | 11.79% (11.65% - 11.94%) |
| ***Race*** | |
| White alone | 75% (74.75% - 75.24%) |
| Black alone | 12% (11.72% - 12.13) |
| Native | 0.1% (0.1% - 0.13%) |
| Asian alone | 5.67% (5.60% - 5.80%) |
| Pacific Islander | 0.17% (0.15% - 0.18%) |
| Some other race alone | 4.34% (4.23% - 4.45%) |
| Mixed race | 2.11% (2.04% - 2.18%) |
| ***Gender*** | |
| Male | 48.45% (48.19% - 48.71%) |
| Female | 51.55% (51.29% - 51.81%) |
| ***Citizenship Status*** | |
| Born in the U.S. | 81.63% (81.42% -81.83%) |
| Born in U.S. territories | 0.68% (0.64%- 0.73%) |
| Born abroad of U.S. parents | 1% (0.09% - 1.01%) |
| U.S. citizen by naturalisation | 8.14% (8.01%-8.28% |
| Not a U.S. citizen | 8.58% (8.41% -8.74%) |

Source: 2015 ACS PUMS
n=234,846
Survey weighted percentages

## 8.3 Results of the OLS models

Next, I will present the results of the analysis. This is done firstly by presenting the results of three OLS regressions that analyse the relationship between the factors of (dis)advantage and the sociodemographic characteristics. Consistent with the argument that intersectionality provides a better lens to understanding disadvantage, an intersectional analysis of multidimensional (dis)advantage follows in which I explore the intersection between race and gender. Therefore, section 8.3.2 presents three additional models (for each dimension of (dis)advantage) that are focused exclusively on intersectionality.

### 8.3.1 The relationship between sociodemographic characteristics and (dis)advantage

To begin exploring the relationship between individual characteristics and the dimensions of (dis)advantage, I looked at the average score on the factors of (dis)advantage for each individual characteristic. These scores - along with their respective 95% confidence intervals - are presented in Table 8.2. The confidence intervals presented in the table indicate whether the (dis)advantage average score for each sociodemographic characteristic is likely to be found in the wider population with 95% confidence. If the confidence intervals crossed zero, it is statistically unlikely that the associated (dis)advantage factor mean will be found in the wider population (Rao and Monroe, 1989).

## Table 8.2: Weighted average score for factors of (dis)advantage by sociodemographic characteristics

| | Labour Force Participation | | | Economic Security | | | Marriage as a Social resource | | |
|---|---|---|---|---|---|---|---|---|---|
| | Mean | S.E. | C.I. | Mean | S.E. | C.I. | Mean | S.E. | C.I. |
| **Age** | | | | | | | | | |
| 65 and up | -0.93 | .004 | (-0.94,-0.93) | 0.06 | .006 | (0.05, 0.07) | -0.02 | .007 | (-0.04, -0.02) |
| 55-64 | 0.05 | .006 | (0.04, 0.06) | 0.16 | .007 | (0.15, 0.18) | -0.01 | .01 | (-0.03, 0.01) |
| 45-54 | 0.39 | .005 | (0.33, 0.39) | 0.12 | .009 | (0.10, 0.13) | 0.01 | .011 | (-0.02, 0.03) |
| 35-44 | 0.43 | .005 | (0.42, 0.44) | -0.11 | .008 | (-0.13, -0.10) | 0.09 | .01 | (0.07, 0.10) |
| 25-34 | 0.41 | .005 | (0.40, 0.42) | -0.35 | .010 | (-0.37, -0.33) | -0.17 | .009 | (-0.19, -0.15) |
| 18-24 | 0.10 | .008 | (0.09, 0.12) | -0.79 | .009 | (-0.81, -0.77) | -0.53 | .005 | (-0.54, -0.52) |
| ***Race*** | | | | | | | | | |
| White | 0.05 | .003 | (0.04, 0.05) | 0.03 | .004 | (0.02, 0.04) | -0.06 | .004 | (-0.06, -0.05) |
| Black | 0.02 | .007 | (0.01, 0.04) | -0.59 | .011 | (-0.61, -0.57) | -0.44 | .011 | (-0.47, -0.42) |
| Native | -0.10 | 0.03 | (-0.16,-0.04) | -0.84 | .032 | (-0.90,-0.78) | -0.30 | .038 | (-0.37, -0.23) |
| Asian | 0.08 | .009 | (0.07, 0.1) | 0.02 | .015 | (-0.01, 0.05) | 0.31 | .014 | (0.28, 0.34) |
| Pacific Islander | 0.10 | .064 | (-0.03, 0.22) | -0.67 | .085 | (-0.85,-0.50) | -0.13 | .083 | (-0.29, 0.03) |
| Some other race alone | 0.19 | .013 | (0.17, 0.22) | -1.23 | .018 | (-1.26,-1.19) | 0.05 | .021 | (0.01, 0.09) |
| Mixed race | 0.14 | .017 | (0.1, 0.17) | -0.35 | .025 | (-0.40, -0.30) | -0.34 | .029 | (-0.39, -0.28) |
| ***Gender*** | | | | | | | | | |
| Male | 0.19 | .003 | (0.18, 0.19) | -0.04 | .005 | (-0.05, -0.03) | -.026 | .005 | (-0.04, -0.02) |
| Female | -0.07 | .004 | (-0.08,-0.06) | -0.19 | .004 | (-0.20, -0.18) | -0.14 | .005 | (-0.15, -0.13) |
| ***Citizenship*** | | | | | | | | | |
| Born in the U.S. | 0.04 | .002 | (0.04, 0.05) | 0.01 | .003 | (0.001, 0.02) | -.158 | 0.004 | (-0.17, -0.15) |
| Born in U.S. territories | -0.14 | .036 | (-0.21,-0.07) | -0.82 | .044 | (-0.91, -0.74) | -0.29 | .051 | (-0.39, -0.19) |
| Born abroad of U.S. parents | 0.23 | .024 | (0.18, 0.28) | 0.01 | .032 | (-.05, 0.07) | -0.05 | 0.04 | (-0.13, 0.03) |
| Naturalised citizen | 0.07 | .011 | (0.05, 0.1) | -0.14 | .014 | (-0.16, -0.11) | 0.23 | .013 | (0.20, 0.26) |
| Not a U.S. citizen | 0.13 | .008 | (0.12, 0.15) | -1.26 | .013 | (-1.23, -1.18) | 0.33 | .012 | (0.31, 0.35) |

Source: 2015 ACS PUMS
N = 234,846
Survey weights applied; S.E. denotes standard error; C.I. denotes the confidence interval

Recall that the average score for each factor is 0, as recognised in Table 7.4 (section 7.5). In initial inspections of Table 8.1, we see that three groups in the sample have below-average advantage in every dimension: 1) women, 2) Native individuals, and 3) those individuals born in U.S. territories.[73] The respective confidence intervals for each of these groups across the factors of (dis)advantage do not cross zero. Therefore, we can be 95% confident that the respective weighted mean can be found in the population and is between the values of the respective confidence interval. No other group has a below-average score in each factor. On the other side of the spectrum, Table 8.2 reveals that only Asians and individuals between the ages of 45 and 54 have an above-average mean in each of the factors of (dis)advantage. More in-depth inspections reveal that for Asian individuals, the confidence interval for 'economic security' crosses zero. This indicates that we cannot say with a 95% confidence that the population mean for Asians can be found within the confidence interval (Rao and Monroe, 1989). The same is revealed for individuals between the ages of 45 and 54 in the 'marriage as a social resource' factor. There were no other groups in this analysis that had consistently above average means in each dimension of (dis)advantage.

What the means in Table 8.2 do not show is whether each category within a sociodemographic characteristic is significantly different from the other categories. For instance, from the table, we cannot tell if Black individuals are significantly different from White individuals in each factor. To examine if there are statistically significant differences, I employ OLS regressions for reasons specified in Chapter 6 (section 6.3). Table 8.3 presents the results of these models, with the factors 'labour force participation,' 'economic security,' and 'marriage as a social resource' as the

[73] Note that this table does not explore the intersectional characteristics. For instance, it does not look at the means for Black women and White women, nor does it look at minority women who are not U.S. citizens. It is, instead, an average for all women.

outcome variable. Age, race, gender, and citizenship status are inputted as independent variables in each model.[74]

In Chapter 6 (section 6.3.1), I acknowledged two goodness of fit statistics that can be used to address the adequacy of the OLS models: 1) the R-square and 2) the F-test. The F-test cannot be computed with the *svyset* commands in Stata. Consequently, I used a Wald statistic to examine the fit on the individual independent variables in the model. The output in Table 8.3 shows that the individual characteristics explain 26% of the variation in the 'labour force participation' factor, 16.3% in 'economic security,' and 3.8% in the 'marriage as a social resource' factor. The Wald test for each of the models was significant, x(16), p<0.001.[75] This indicates that each sociodemographic characteristic included in the model is significant and should remain in each model predicting each dimension of (dis)advantage. Interestingly, I find that the sociodemographic characteristics explain nearly six times the variation in labour force participation than it does in 'marriage as a social resource.'

---

[74] It is recognised that additional independent variables may be beneficial in understanding the different dimensions of (dis)advantage. For instance, educational attainment may be associated with labour force participation. As educational attainment is included in the 'economic security' factor, it is not analysed as an independent variable. It could be included in a separate OLS model where 'labour force participation' is the outcome, but it is the objective of this chapter to explore if and how characteristics beyond an individual's control are associated with each of the factors of (dis)advantage and to explore that variation, should it exist. Exploring any association with additional independent variables is beyond the focus of this chapter.

[75] For 'labour force participation,' $X^2(16) = 155{,}553.80$. For 'economic security,' $X^2(16) = 41{,}230.84$. For 'marriage as a social resource, = $X^2(16) = 21{,}282.97$.

## Table 8.3: OLS regression models of the factors of (dis)advantage on individual sociodemographic characteristics

| | Labour Force Participation Coef. (Std. Error) | Economic Security Coef. (Std. Error) | Marriage: a social resource Coef. (Std. Error) |
|---|---|---|---|
| Constant | -.801*** | 0.283*** | 0.022*** |
| | (0.005) | (0.007) | (0.008) |
| ***Age*** | | | |
| 65 and up (ref) | | | |
| 55-64 | 0.979*** | 0.134*** | 0.018 |
| | (0.006) | (0.007) | (0.011) |
| 45-54 | 1.321*** | 0.143*** | 0.019 |
| | (0.006) | (0.01) | (0.012) |
| 35-44 | 1.371*** | -0.011 | 0.081*** |
| | (0.006) | (0.009) | (0.011) |
| 25-34 | 1.353*** | -0.279*** | -0.155*** |
| | (0.007) | (0.011) | (0.011) |
| 18-24 | 1.044*** | -0.768*** | -0.477*** |
| | (0.009) | (0.012) | (0.008) |
| ***Race*** | | | |
| White alone (ref) | | | |
| Black alone | -0.116*** | -0.571*** | -0.364*** |
| | (0.008) | (0.01) | (0.012) |
| Native | -0.239*** | -0.827*** | -0.213*** |
| | (0.031) | (0.033) | (0.036) |
| Asian alone | -0.072*** | 0.468*** | 0.145*** |
| | (0.011) | (0.017) | (0.015) |
| Pacific Islander | -0.067 | -0.379*** | -0.048 |
| | (0.049) | (0.082) | (0.083) |
| Other race alone | -0.015 | -0.704*** | -0.037 |
| | (0.012) | (0.018) | (0.021) |
| Mixed race | -0.057*** | -0.202*** | -0.228*** |
| | (0.051) | (0.024) | (0.03) |
| ***Gender*** | | | |
| Male (ref) | | | |
| Female | -0.217*** | -0.163*** | -0.117*** |
| | (0.004) | (0.006) | (0.007) |
| ***Citizenship status*** | | | |
| Born in the U.S. (ref) | | | |
| Born in U.S. territories | -0.197*** | -0.745*** | -0.16** |
| | (0.03) | (0.047) | (0.053) |
| Born abroad of U.S. parents | 0.012 | -0.019 | 0.101** |
| | (0.023) | (0.033) | (0.39) |
| U.S. citizen by naturalisation | 0.081*** | -0.346*** | 0.285*** |
| | (0.01) | (0.016) | (0.014) |
| Not a U.S. citizen | -0.103*** | -1.2*** | 0.435*** |
| | (0.008) | (0.016) | (0.014) |
| R-squared | 0.262 | 0.163 | 0.038 |
| Number of observations | 234,846 | 234,846 | 234,846 |

Source: 2015 American Community Survey Public Use Microdata Files
Survey weights applied; ref indicates reference category
Standard errors in parentheses
*p<0.05, **p<0.01, *** p<0.001

Because I am concerned with the extent of the relationship between the individual sociodemographic variables[76] and the factors of disadvantage, I examine the independent variables and compare the results across the three models in lieu of reporting the results of each model singularly. I start with age.

In the 'labour force participation' model, there is a statistically significant (*p<0.001*) difference between each age group and the reference category (individuals aged 65 and over). Holding the other independent variables constant, the respective coefficients are also higher than the reference category. These differences in the coefficients follow a life cycle pattern of labour force participation recognised in the literature. For instance, Abraham (2015) recognises that young adults who may still be in school work less; labour force participation increases during the middle years of life and falls again as one nears retirement. The peak average score in this model is for individuals aged between 35 and 44. Interestingly, this statistically significant difference between individuals aged between 35 and 44 and the reference group is not found in the 'economic security' model. In fact, there is not a statistically significant difference. All of the other age groups are statistically different from the reference group. Taking into consideration the components of the economic security factor, we could expect that younger individuals would have less advantage than the reference group of individuals. For instance, older Americans typically have access to government-provided health insurance and often have access to Social Security income, which increases their total income levels. Health insurance coverage and total individual income, both have a positive relationship with the economic security factor (presented in Table 7.3). In the marriage as a social resource dimension, individuals aged between 18 and 34, experience less advantage, on average, compared to individuals aged 65 sand up. For individuals aged 35-44, there is a positive difference in the 'marriage as a social resource'

[76] For these variables, I tested for multicollinearity in the models as mentioned in Chapter 6. Multicollinearity can be tested via the variance inflation factor (VIF), which reports how much of the variance of the estimated coefficients for the independent variables (here, the sociodemographic characteristics) increases due to collinear variables (Craney and Surles, 2002). Mansfield and Helms (1982) acknowledge that if the VIFs are not substantially larger than 1.0, then multicollinearity is not an issue. For these variables, each VIF was between 1.0 and 1.2, suggesting that multicollinearity is not present in the models.

factor (*b* = 0.081, *p*<.001), compared to the reference group. This is the only age group that had a statistically significant positive difference in this factor to our reference group, holding all the other sociodemographic characteristics in the model.

Examining sex, the results in Table 8.3 shows that being a female is associated with less advantage in each factor compared to men. The most considerable average difference in advantage between men and women is found in the 'labour force participation' factor (*b* = -0.22, *p*<0.001). This difference is particularly impressive given that the most told story in recent decades has been that of the increased female labour force participation (Juhn and Potter, 2006; Treas, 1987). There could be a couple of reasons for this difference. Firstly, in the data, the most represented category is women over the age of 65. These women are probably retired and not actively seeking work. Secondly, it could be that women spend less time in paid work than men do in unpaid work (Sayer *et al.*, 2009). Note that the ACS PUMS does not include information on unpaid work, such as running a household or providing childcare, so I am unable to capture these experiences that tend to be taken up mostly by women (Cherlin, 2004). Thirdly, my interpretation of this factor expands current discussion of labour force participation to include actual participation in the labour force but also the ability to participate that includes a recognition of spatial inequality (section 7.4.1), like travel time to work in which there are substantial gendered differences (Schwanen and Dijst, 2002; Turner and Niemeier, 1997). It is essential to discuss it as it relates to the debates on gender - and racial inequalities - in the labour market (Elliott and Joyce, 2004). For gender, a common argument is that societal and cultural expectations associated with marriage and motherhood encourage women to choose jobs closer to home, which shrinks their opportunities in the labour market and subsequently their wages (Gangl and Ziefle, 2009). The differences between men and women exist in the other dimensions, though they are not as large as in 'labour force participation.' Interestingly, the smallest difference is in the 'marriage as a social resource' factor. Monin and Clark (2011) suggest that men receive more benefits inside of marriage than women do, which could serve as an explanation for the differences in advantage in the 'marriage as a social resource' factor (*b* = -0.12, *p*<0.001). This is a particularly important finding as the items included in this variable do not contain any social connections outside of the union between two people, whether it be marriage, divorce, or an unmarried partnership. This does highlight a characteristic of this data, in that the most counted individual in this analysis has the characteristic of being a White woman aged 65 and up. This could reflect some of the issues of widowhood that was addressed in section 7.3.1 that continuously loads on factor by itself. The results here highlighted some important differences in (dis)advantage based on gender.

Regarding race across the factors of (dis)advantage, the model reports that those who are Black ($b$ = -0.12, $p$<0.001), Native ($b$ = -0.24, $p$<0.001), Asian ($b$ = -0.07, $p$<0.001) and of Mixed race ($b$ = 0.06, $p$<0.001) have less advantage in the 'labour force participation' than White individuals, on average and holding the other characteristics under analysis constant. Despite the lack of advantage in the labour force participation factor, Asians, compared to Whites, have significantly more advantage in 'economic security' (b =0.47, p<0.001) and 'marriage as a social resource' ($b$=0.15, $p$<0.001). The results here contradict the findings of Segal *et al.* (2002), who argue that Asians in the United States have lower incomes (when controlling for education) and higher poverty rates than Whites. It is worth noting that income poverty, educational status, and total individual income are included in the 'economic security' factor, where the results show that Asians in the United States, on average, have higher advantage than Whites, holding the other variables in the model constant. Though the results here contradict their findings, Segal *et al.* (2002) do rightly acknowledge that race in the United States continues to be significant for the well-being of minority groups, and for particularly Asian groups, making them disadvantaged in American society. My analysis supports a portion of their argument but does not extend to Asians in the 'economic security' and 'marriage as a social resource' dimensions. My results suggest that Asians are relatively more advantaged than Whites in two of the three factors. The differences in findings could be so because 'Asian' contains a heterogeneous group of people from various cultures and many different countries. Segal *et al.*

(2002) recognise this heterogeneity, but they group Pacific Islanders with other Asians.[77] They also utilise information from the Current Population Reports, which are derived from the Current Population Survey. I discussed in Chapter 4 (section 4.5) some of the limitations of the Current Population Survey. The results of my analysis, using the ACS PUMS, uncover less advantage on average for other minority groups, compared to White Americans in multiple dimensions. That is not always the case for the Asians in this data set. For instance, according to the analysis, in addition to being less advantaged in the 'labour force participation' factor, Black respondents have the highest average difference in 'economic security' ($b$ = -0.83, $p$<0.001)[78], from Whites. Black is also associated with the highest difference, on average, ($b$ = -0.36, $p$<0.001) in the marriage as a social resource factor, compared to Whites, holding our characteristics in the model constant. The results show that there is a link between multiple dimensions of disadvantage for Black Americans. The analysis lends weight to an argument made by Bhopal (2018), who argues that Black and other minority ethnic groups in neoliberal societies, such as the United States, still face inequalities not experienced by Whites. This can be seen particularly in labour force participation, as residential segregation by race has placed minorities further away from job opportunities, inhibiting their ability to find suitable employment opportunities (Elliott and Joyce, 2004; Kain, 1968; Mooney, 1969). The results presented here extend this argument further in the United States context because there is first a recognition that there are multiple (dis)advantages beyond income. In the second instance, different segments of the population may be more advantaged in one dimension and less so in another. The analysis finds that Asians in the United States do not face more disadvantage than Whites, on average, in a society marked by 'white privilege' (Bhopal, 2018), but enjoy more advantage in 'economic security' and 'marriage as a social resource.'

[77] They highlight a statistical directive from 1977 from the United States' Office of Budget and Management that sought to standardise racial and ethnic group membership, in which Asians (those with lineage from Asia) were lumped with Pacific Islanders (Segal *et al.*, 2002).

[78] Individuals who identify as other race alone had the second highest difference in economic security compared to Whites ($b$ = -0.70, $p$<0.001).

Lastly, the results for citizenship status indicate that individuals born in American territories have significantly less advantage, on average, in each of the factors than do individuals born in the United States. The largest difference between the groups exists in the 'economic security' factor ($b$ = -0.75, $p$<0.001). There is also a statistically significant difference in 'economic security' for individuals who are not a U.S. citizen and those born in the United States ($b = -1.2$, $p<0.001$). Interestingly, for those who are not a U.S. citizen there is a significantly positive difference ($b = 0.44$, $p<0.001$) in the 'marriage as a social resource' factor compared to individuals born in the United States. This lends weight to an argument made by Valdez *et al*. (2013) that non-American citizens stay in the United States because of familial and community factors, indicating that they have substantial social resources. This could also be true for those who are American citizens by naturalisation, who also have a significantly positive difference ($b = 0.29$, $p<0.001$) in the 'marriage as a social resource' factor compared to those who are born in the United States. This is supported in the literature, which suggests that individuals who become American citizens via naturalisation often have secure social connections via various group memberships and a strong community environment (Gubernskaya *et al.*, 2013; Logan *et al.*, 2012).

From this analysis, the benefit of examining (dis)advantage multidimensionally is highlighted. Taken together, these analyses provide preliminary evidence that, as Kabeer (2000) acknowledges, different forms of disadvantage have the potential to give rise to different kinds of disadvantaged groups. It also highlights that with other variables being held constant, Black Americans and women are consistently and significantly less disadvantaged than Whites and men, respectively, on average. In the next section, I conduct three additional OLS regressions, but with a focus on the intersection between race and gender.

### 8.2.3 An intersectional analysis of multidimensional disadvantage

The aim of this section is to extend the previous analysis by exploring the relationship between the factors of (dis)advantage and sociodemographic characteristics via an intersectional framework.

There is a lack of quantitative empirical work on intersectionality; the complexity of measuring it has been well-recognised in the literature (Bowleg, 2008; Dubrow, 2008; McCall, 2005). As such, this thesis contributes to the literature on poverty and disadvantage, by working through the complexity and quantitatively analysing intersectionality. The intersectional focus in this analysis is the one that exists between race and gender because this is the intersection of identity first used by Crenshaw (1991) who articulated intersectionality (recognised in section 2.5 and 6.3). Similar to section 8.3.1, I will first present the results of the weighted means for women across race. Then I will address the results of the OLS models. Section 6.3 highlighted that the multiplicative approach is the ideal approach to empirically analyse intersectionality in OLS models because it embodies a core tenet of intersectionality that emphasises that the influence of a given demographic characteristic on a social outcome is conditional on the intersection of characteristics (Dubrow, 2008).

Table 8.4 presents the weighted means of the three factors of (dis)advantage, 'labour force participation,' 'economic security,' and 'marriage as a social resource,' for the intersecting identities of interest. For each group of means, the 95% confidence intervals are presented. As addressed in the previous section, the confidence intervals presented in the table indicate whether each category's mean is likely to be found in the broader population with 95% confidence (Rao and Monroe, 1989). What can be gathered from Table 8.4 is that there is 95% confidence that the mean for Asian women's labour force participation is between -0.097 and -0.046. However, the same cannot be said for Black women alone or Mixed Race women. For 'economic security,' only the mean presented for Asian women is associated with a confidence interval that crosses zero, suggesting that we cannot be 95% confident that the mean presented in Table 8.4 would be found in the population. For 'marriage as a source resource,' it is not statistically likely with 95% confidence that the means presented in the table for Pacific Islander and Other Race women will be found in the wider population for these groups.

## Table 8.4: Weighted means of (dis)advantage factors for intersectional sociodemographic characteristics

| Intersecting identities (race and sex) | % | Mean of 'labour force participation' (confidence intervals) | Mean of 'economic security' (confidence intervals) | Mean of 'marriage as a social resource (confidence intervals) |
|---|---|---|---|---|
| Gender and race | | | | |
| White Women (alone) | 38.26% | -.084 (-0.09, -.08) | -0.048 (-0.06, -0.04) | -.102 (-0.12, -0.09) |
| Black women (alone) | 6.55% | -.011 (-0.03, 0.01) | -.638 (-0.67, -0.61) | -.543 (-0.57, -0.52) |
| Native women | 0.41% | -.165 (-0.24, -0.09) | -.866 (-0.95, -0.78) | -.324 (-0.43, -0.22) |
| Asian women (alone) | 3.02% | -.072 (-.1, -.05) | -.019 (-.06, .02) | .292 (0.26, 0.33) |
| Pacific Islander women | 0.09% | -.14 (-.31, .03) | -.888 (-1.11, -.67) | -.068 (-0.32, 0.18) |
| Other race women | 2.17% | -.023 (-0.06, 0.01) | -1.278 (-1.33, -1.23) | -.001 (-0.06, 0.06) |
| Mixed race women | 1.05% | .034 (-0.01, 0.08) | -.395 (-0.45, -0.34) | -.350 (-0.43, -0.27) |

Source: 2015 ACS PUMS, survey weights used
n=234,846
Social exclusion factors, true mean = 0
Group means based on weighted data

Next, the results of the OLS regressions are presented in Table 8.5. It is important to note that when one employs the multiplicative approach, it is necessary to include the categories that make up the intersection along with the interaction (Dubrow, 2008; McCall, 2005). Therefore, the independent variables in the models are race, gender, and the interactions between race and gender.

Two goodness of fit statistics were calculated. Firstly, the Wald test for the overall model fit at $X^2$ (13) was statistically significant for each of the models ($p<0.001$).[79] Secondly, the values of R-square across the models indicate that the independent variables explain 1.85% of the variation in 'labour force participation,' 6.21% in 'economic security,' and 1.53% in 'marriage as a social resource.' This suggests that race, gender, and their interaction explain three times more variation in 'economic security,' compared to 'marriage as a social resource.' The variation uncovered here indicates that race and gender are important in understanding the variation of disadvantage in the United States, particularly in 'economic security.' The importance of economic security for individuals with the identities under investigation has been highlighted by Morris and Deprez (2014). They acknowledge

[79] For 'labour force participation,' $X^2(13)$ = 8,609.66. For 'economic security,' $X^2(13)$ = 3,563.87. For 'marriage as a social resource, = $X^2(13)$ =2,981.57

key social and policy issues. Firstly, regulations in the United States influence, often negatively, health care services for women, and neglect addressing the high rates of income poverty faced by women, particularly women of colour (Morris and Deprez, 2014). In addition, affirmative action policies that affect university admissions for women of colour, are under review, which can hinder their educational attainment aspirations (Morris and Deprez, 2014). Finally, if women work year-round, compared to men, they still have lower levels of income (Morris and Deprez, 2014). Each issue that Morris and Deprez (2014) address are included in the 'economic security' factor of (dis)advantage in this analysis. This highlights and supports their argument that women, in particular minority women, are at continued risk of economic insecurity in the United States.

**Table 8.5: OLS regression models of the factors of (dis)advantage on intersectional characteristics**

| | Labour Force Participation Coef./Std. err | Economic Security Coef./Std. err | Marriage as a Social Resource Coef./Std. err |
|---|---|---|---|
| Constant | 0.183*** | 0.112*** | -0.005*** |
| | (.004) | (.005) | (.006) |
| ***Race*** | | | |
| White alone (ref) | | | |
| Black alone | -0.117*** | -0.635*** | -0.316*** |
| | (.013) | (.016) | (.019) |
| Native | -0.211*** | -0.924*** | -0.266** |
| | (.038) | (.05) | (.06) |
| Asian alone | 0.074*** | -0.058* | 0.333*** |
| | (.014) | (.024) | (.021) |
| Pacific Islander | 0.179* | -0.538*** | -0.191 |
| | (.073) | (.123) | (.111) |
| Other race alone | 0.228*** | -1.284*** | 0.105*** |
| | (.015) | (.027) | (.025) |
| Mixed race | -0.056*** | -0.412*** | -0.317*** |
| | (.024) | (.04) | (.037) |
| ***Gender*** | | | |
| Male (ref) | | | |
| Female | -0.267*** | -0.159*** | -0.101*** |
| | (.006) | (.007) | (.009) |
| ***Interaction: race and gender*** | | | |
| White women (ref) | | | |
| Black women alone | 0.19*** | 0.045* | -0.122*** |
| | (.017) | (.022) | (.024) |
| Native women | 0.129* | 0.106 | 0.047 |
| | (.051) | (.064) | (.086) |
| Asian women | -0.061** | 0.086** | 0.064* |
| | (.02) | (.03) | (.027) |
| Pacific Islander women | -0.235* | -0.302 | 0.228 |
| | (0.11) | (.165) | (.173) |
| Other race alone women | -0.167*** | 0.054 | 0.0002 |
| | (.023) | (.035) | (.042) |
| Mixed race women | 0.063 | 0.065 | .071 |
| | (.035) | (.051) | (.052) |
| R-squared | 0.0185 | 0.0621 | 0.0153 |
| No. of observations | 234,846 | 234,846 | 234,846 |

Source: 2015 American Community Survey Public Use Microdata Files
Survey weights applied
Standard errors in parentheses
*p<0.05, **p<0.01, *** p<.001

Despite the varying level of variance explained by the independent variables across the models, there are some consistent patterns. The relationships uncovered between singular categories, race and gender in Table 8.2 are consistent with the results uncovered in Table 8.4. For instance, women are less advantaged in each category compared to men. For the main effects for race, I find that the relationships in Table 8.4 are similar to the relationships uncovered in Table 8.2 in 'labour force participation' and 'marriage as a social resource.' There is a difference between the two models for 'economic security.' In the intersectional model (Table 8.4), all of the racial minorities are less advantaged in the 'economic security' factor, compared to whites. The largest difference is between other race alone and White ($b = -0.704$, $p<0.001$) in the non-intersectional model and between Native and White in the intersectional model ($b = 0.924$, $p<0.001$). While important to recognise these similarities and differences across the various OLS models, the focus in this section is on the intersectional relationships (gender and race) with the multiple factors of disadvantage.

Looking at the coefficients for the intersectional characteristics of interest, race and gender, I do not find a lot of consistent statistically significant differences for women by race across the three dimensions of (dis)advantage. There are some results I find particularly interesting. Firstly, I find that most of the significant difference between minority women and White women are in the 'labour force participation' factor. Secondly, when I compare the results for Mixed Race women to White women, I do not find any statistically significant difference in any of the factors of (dis)advantage. This suggests that the incidence or level of (dis)advantage for White and Mixed Race women is not notably different. For the other women, I do find differences. Pacific Islander women and Native women have significantly ($p<0.05$) less advantage in the 'labour force participation' factor compared to White women. Three groups of women have on average significantly more advantage in this factor, compared to White women: 1) Black ($p<0.001$), 2) Asian ($p<0.01$), and 3) other race women ($p<0.001$). Interestingly, in the other two factors, only Black women and Asian women have significant differences compared to White women. All of the other groups of women have no average differences in 'economic security' and 'marriage as a social resource.'

Crenshaw (1991) articulated the intersectionality framework, particularly with Black women in mind as their intersection of identities made them alienated by anti-racist and pro-feminist movements. For me, it is important to continue with her original focus and interpret the story of Black women in the United States. There remains a gap in the literature that highlights their experiences and challenges. I use the intersectional approach to analyse the Black woman's experience of multidimensional disadvantage because of the complex social context in which they live (Johnson and Loscocco, 2015). From this analysis and other research in the literature, it has been highlighted that women, compared to men, and racial minorities, compared to Whites, have higher instances of disadvantage (Morris and Deprez, 2014; Proctor *et al.*, 2016). However, when compared to White women in the 'labour force participation' factor, Black women, on average, have significantly ($p<0.001$) more advantage. Because most of an individual's income is derived from employment, we would expect to see a similar pattern between these two groups in the 'economic security' factor, particularly as income is a component of this factor. This is not the case. Black women may have more advantage in the 'labour force participation' factor, but they are less advantaged in 'economic security,' much less than White women.[80] The results highlight two issues. Firstly, the findings lend weight to some European literature that suggests that inclusion - or a reduction of social exclusion- should not rely solely on increasing the labour force participation rate (Leach *et al*., 2010). Secondly, these findings point to a need to address issues of working poverty and underemployment (Brady *et al.*, 2013; Crettaz and Bonoli, 2010; Kalleberg, 2012; Kalleberg *et al.*, 2000). In a report conducted by DuMonthier *et al*. (2017) for The Institute for Women's Policy Research, it was found that Black women have some of the highest labour force participation rates in the United States. However, their income lags behind men and women of other races and they have low rates of health insurance among the nonelderly (DuMonthier *et al*., 2017). Both income and health insurance are components of the 'economic security' factor, in which income and having health insurance has a positive

[80] The OLS regression for economic security predicts that for Black women the score, considering the other independent variables in the model will be -0.637 and -0.047 for White women.

association with this factor and lower levels of income and no health insurance are issues that seem to be faced proportionally higher by Black women in the United States (DuMonthier *et al*., 2017).

The relationship between Black women and multidimensional disadvantage is particularly interesting compared to the findings for Asian women. I make this comparison because Black and Asian women were the only two groups of women to have a statistically significant difference in each factor of (dis)advantage compared to White women. Firstly, holding all the other variables constant, Black women have more advantage in 'labour force participation,' more than White and Asian women. Despite less advantage in the 'labour force participation' factor, Asian women enjoy more advantage, on average, in both the 'economic security' and the 'marriage as a social resource' factors. Labour force participation is often promoted as a means to economic security and prosperity (Besley and Coate, 1992; Yeo and Moore, 2003). However, this analysis suggests that for Black and Asian women in the United States, this suggestion does not seem to hold true. Black women have less advantage in the 'economic security' factor despite the advantage in 'labour force participation.' Previously in this chapter, I pondered the reasons Black women may have a relatively lower advantage in 'economic security.'

Secondly, Asian women have a higher advantage in 'marriage as a social resource' compared to White and Black women. A search of the literature reveals some consistent generalisations about Asian Americans. The literature points to Asian women's willingness to engage in interracial marriages (Kitano et al., 1984; Min and Kim, 2009; Qian, 2005), which increases the sex ratio of potential partners and influences the transitions in and out of marriage (Warner et al., 2011) Kitano et al. (1984) recognise that Asian women are more likely than Asian men to marry outside of their race. This type of integration is less available for Black women where racial prejudice lingers and tends to limit this intimate form of integration (Qian, 2005). Race has been cited elsewhere in the literature. For instance, Johnson and Loscocco (2015) acknowledge that marriage for Black women is different for other groups, which in part comes from trying to create intimate ties with men in the context of a racist society while also receiving fewer benefits than what White women tend to receive from marriage. Ridgeway and Kricheli-Katz (2013) suggest that Black women's failure to live up to dominant (White) femininity makes them less marriageable than White and Asian women. Many Black women choose to remain single, reflecting the voluntary nature of self-exclusion. However, Johnson and Loscocco (2015) admit these choices may be constrained by the realities of their positions as subordinates as women and black individuals. It may be that the ability of Asian women to participate in interracial marriages opens them up to the economic benefits of marriage that Black women are not able to enjoy (Hurtado, 1989), thus enabling them to enjoy more advantage in the 'economic security' factor.

The analysis reveals an interesting pattern of (dis)advantage in the United States. In adopting an intersectional framework in my analysis, I found significant relationships that a focus on singular sociodemographic characteristics was unable to uncover. It further highlights that different forms of disadvantage can are experienced differently by different groups of the population (Kabeer, 2000).

## 8.4 Conclusion

The results of this analysis lead to several conclusions. Firstly, the sociodemographic characteristics of age, race, gender, and citizenship status explain 26% of the variation in 'labour force participation,' whereas it explains 4% in 'marriage as a social resource.' This indicates that characteristics outside of an individual's control explain over one-quarter of their participation in the labour market. Secondly, the intersectional analysis reveals that three times more variation in 'economic security' is explained by intersecting characteristics than it does in the other two factors. This lends weight to a consistent argument in this thesis that (dis)advantage is not felt universally and that the multiple forms of (dis)advantage influence different groups of people differently. This highlights that policies in the United States developed to tackle disadvantage should be made in recognition of these differences.

# Chapter 9

# Introducing the state: contextual heterogeneity in (dis)advantage

## 9.1 Introduction

This chapter expands the results of the previous chapters by introducing the American state as a level of analysis utilising multilevel modelling. Primarily, I am examining the extent to which variation in each dimension of (dis)advantage exists across the United States. In exploring the state as a level of analysis, this chapter will answer the following research question:

*Is there variation in multidimensional disadvantage across the United States?*

In Chapter 6, I acknowledged that there is potential to expand the variance components multilevel model used to analyse the research question above. In this chapter, I show that there is significant variation across the United States in each dimension of disadvantage. Subsequently, I extend the variance components multilevel model to assess if that variation persists whilst controlling for individual characteristics, particularly age, race, gender, and citizenship status. These extended models allow me to answer the following sub-questions:

*Does that variation, if any, still persist after controlling for individual characteristics?*

*Does the relationship between individual sociodemographic characteristics and multidimensional disadvantage vary significantly across U.S. states and the District of Columbia?*

## 9.2 Analysis

The multilevel modelling of 234,846 individuals aged 18 and up nested within 51 groups (the 50 U.S. states and Washington, District of Columbia, from now referred to as states) is achieved via the *mixed* command in Stata version 15 statistical software. All parameters were estimated using the full maximum likelihood method for reasons specified in chapter 6 (section 6.4.1).

This chapter has three main objectives, for which the results of twelve multilevel models are presented. The first objective is to determine if there is any variation in multidimensional (dis)advantage at the state level. This objective is reached via the examination of three variance components or null models, one for each dimension of (dis)advantage. The null models estimate the contribution of the state to the variance of the dependent variable and are benchmark models that can be used to compare later models as it includes no independent variables. I recognised in section 6.4.1 that if there is variation found amongst the states, it would be possible to do further analysis. The second and third objectives (discussed next) and the additional research questions stated in section 9.1 reflect this.

The second objective is to test state-level variation in (dis)advantage whilst controlling for individual sociodemographic variables. For that purpose, the modelling strategy is geared toward examining the extent to which variation remains amongst the states for each dimension of (dis)advantage, holding individual characteristics constant. For this purpose, the null models are expanded to random intercepts models and include the individual sociodemographic characteristics. Consistent with the previous chapter, intersectionality is also explored. Therefore, there are two random intercepts models for each dependent variable to explore the non-intersectional (age, race, gender, and citizenship status) and intersectional (gender and race) individual characteristics. All independent variables are operationalised just as they are in the previous chapter.

The last objective is to assess the roles played by individual sociodemographic characteristics by not holding them constant across states in explaining the observed variation in each dimension of (dis)advantage within and across states. This is achieved via the use of random coefficient multilevel models.

This chapter is examining the extent to which variation in (dis)advantage exists within and between states. I do not include any second level (state) predictors to any of the models presented. There are no state-level variables available in the American Community Survey (ACS) Public Use Microdata Sample (PUMS) files. Adding state predictors would require a combination of many different data sets, a task which is not undertaken for this thesis due to time and resource restrictions. It is recognised that the state-level predictors, such as the aggregate percentages of individuals with the various sociodemographic characteristics examined in this study, would aid in understanding the variation of each dimension of (dis)advantage across states and as such, there remains much potential for future research with state-level predictor variables. This is further discussed in the concluding chapter of this thesis (chapter 10, section 10.7). The following sections report the findings and then the chapter is concluded.

## 9.3 Descriptive Results

Table 9.1 presents each state's average score for each dimension of (dis)advantage, along with the respective 95% confidence intervals. The lowest average score in labour force participation is found in West Virginia (average = -0.108) and the highest is in North Dakota (0.254). In the 'economic security' dimension, Mississippi has the lowest average score with a value of -0.34. The highest average score is found in New Hampshire with a value of 0.36. The District of Columbia has the lowest average score in 'marriage as a social resource,' with a value of -0.49. Utah has the highest, with an average score of 0.17. The 'economic security' dimension has the highest range of scores with a value of 0.7. The 'labour force participation' dimension has the lowest range of average scores, 0.362. The range for the 'marriage as a social resource' dimension is quite close to the economic security dimension with a value of 0.66.

## Table 9. 1: Average score in (dis)advantage and confidence intervals by state

| State | Labour force participation | | | Economic security | | | Marriage as a social resource | | |
|---|---|---|---|---|---|---|---|---|---|
| | Mean | 95% CI | | Mean | 95% CI | | Mean | 95% CI | |
| Alabama | -0.06 | -0.10 | -0.01 | -0.24 | -0.30 | -0.19 | -0.08 | -0.15 | -0.01 |
| Alaska | 0.09 | -0.04 | 0.23 | -0.19 | -0.35 | -0.04 | -0.03 | -0.20 | 0.13 |
| Arizona | -0.08 | -0.11 | -0.04 | -0.22 | -0.27 | -0.17 | -0.16 | -0.22 | -0.11 |
| Arkansas | -0.07 | -0.12 | -0.03 | -0.27 | -0.33 | -0.21 | -0.11 | -0.19 | -0.03 |
| California | 0.06 | 0.04 | 0.07 | -0.24 | -0.26 | -0.23 | -0.03 | -0.05 | -0.01 |
| Colorado | 0.16 | 0.13 | 0.20 | 0.15 | 0.11 | 0.19 | -0.13 | -0.19 | -0.07 |
| Connecticut | 0.13 | 0.08 | 0.17 | 0.22 | 0.17 | 0.27 | -0.06 | -0.12 | 0.00 |
| Delaware | 0.03 | -0.06 | 0.12 | -0.03 | -0.15 | 0.09 | -0.17 | -0.30 | -0.04 |
| D.C. | 0.07 | -0.02 | 0.15 | 0.20 | 0.06 | 0.34 | -0.49 | -0.61 | -0.37 |
| Florida | -0.05 | -0.07 | -0.03 | -0.28 | -0.30 | -0.25 | -0.18 | -0.21 | -0.16 |
| Georgia | 0.04 | 0.01 | 0.07 | -0.20 | -0.23 | -0.16 | -0.06 | -0.09 | -0.03 |
| Hawaii | 0.12 | 0.06 | 0.19 | -0.11 | -0.22 | 0.01 | 0.03 | -0.06 | 0.12 |
| Idaho | 0.01 | -0.06 | 0.09 | -0.22 | -0.32 | -0.12 | -0.05 | -0.16 | 0.06 |
| Illinois | 0.09 | 0.07 | 0.11 | -0.05 | -0.08 | -0.02 | -0.07 | -0.11 | -0.03 |
| Indiana | 0.07 | 0.03 | 0.11 | -0.12 | -0.16 | -0.07 | -0.14 | -0.19 | -0.10 |
| Iowa | 0.15 | 0.10 | 0.19 | -0.02 | -0.08 | 0.04 | -0.09 | -0.18 | 0.01 |
| Kansas | 0.11 | 0.07 | 0.16 | 0.05 | -0.02 | 0.13 | -0.13 | -0.23 | -0.04 |
| Kentucky | -0.03 | -0.08 | 0.01 | -0.19 | -0.24 | -0.14 | -0.12 | -0.17 | -0.06 |
| Louisiana | -0.02 | -0.06 | 0.03 | -0.33 | -0.38 | -0.27 | -0.19 | -0.25 | -0.13 |
| Maine | 0.03 | -0.05 | 0.10 | -0.01 | -0.12 | 0.10 | -0.10 | -0.23 | 0.02 |
| Maryland | 0.15 | 0.12 | 0.19 | 0.24 | 0.20 | 0.29 | -0.07 | -0.12 | -0.02 |
| Massachusetts | 0.12 | 0.09 | 0.15 | 0.25 | 0.21 | 0.29 | -0.14 | -0.19 | -0.09 |
| Michigan | 0.02 | 0.00 | 0.05 | -0.08 | -0.12 | -0.05 | -0.13 | -0.17 | -0.08 |
| Minnesota | 0.18 | 0.14 | 0.22 | 0.23 | 0.18 | 0.28 | -0.01 | -0.07 | 0.05 |
| Mississippi | -0.03 | -0.09 | 0.02 | -0.37 | -0.43 | -0.30 | -0.14 | -0.22 | -0.06 |
| Missouri | 0.03 | -0.01 | 0.07 | -0.11 | -0.15 | -0.06 | -0.11 | -0.16 | -0.06 |
| Montana | 0.01 | -0.08 | 0.10 | -0.07 | -0.20 | 0.06 | -0.12 | -0.29 | 0.04 |
| Nebraska | 0.24 | 0.18 | 0.30 | 0.02 | -0.07 | 0.10 | -0.02 | -0.13 | 0.09 |
| Nevada | 0.05 | 0.00 | 0.09 | -0.30 | -0.36 | -0.23 | -0.20 | -0.28 | -0.12 |
| New Hampshire | 0.22 | 0.15 | 0.29 | 0.36 | 0.28 | 0.44 | -0.05 | -0.18 | 0.08 |
| New Jersey | 0.09 | 0.06 | 0.12 | 0.12 | 0.08 | 0.16 | 0.01 | -0.03 | 0.04 |
| New Mexico | -0.05 | -0.12 | 0.02 | -0.31 | -0.40 | -0.22 | -0.18 | -0.28 | -0.08 |
| New York | 0.04 | 0.03 | 0.06 | -0.10 | -0.13 | -0.07 | -0.09 | -0.11 | -0.06 |
| North Carolina | 0.01 | -0.01 | 0.04 | -0.22 | -0.26 | -0.18 | -0.10 | -0.14 | -0.06 |
| North Dakota | 0.25 | 0.15 | 0.36 | 0.17 | 0.06 | 0.28 | -0.10 | -0.23 | 0.04 |
| Ohio | 0.06 | 0.04 | 0.08 | -0.04 | -0.07 | 0.00 | -0.11 | -0.14 | -0.07 |
| Oklahoma | 0.03 | -0.02 | 0.08 | -0.23 | -0.29 | -0.17 | -0.10 | -0.17 | -0.03 |
| Oregon | 0.02 | -0.03 | 0.06 | -0.22 | -0.29 | -0.15 | -0.16 | -0.23 | -0.09 |
| Pennsylvania | 0.04 | 0.01 | 0.07 | 0.01 | -0.03 | 0.05 | -0.10 | -0.13 | -0.07 |

*Continued on next page*

| State | Labour force participation | | | Economic security | | | Marriage as a social resource | | |
|---|---|---|---|---|---|---|---|---|---|
| | Mean | 95% CI | | Mean | 95% CI | | Mean | 95% CI | |
| Rhode Island | 0.14 | 0.06 | 0.22 | 0.11 | 0.01 | 0.21 | -0.17 | -0.28 | -0.06 |
| South Carolina | -0.35 | -0.07 | 0.001 | -0.24 | -0.29 | -0.19 | -0.06 | -0.13 | 0.001 |
| South Dakota | 0.20 | 0.10 | 0.30 | 0.15 | 0.04 | 0.26 | 0.08 | -0.07 | 0.22 |
| Tennessee | -0.01 | -0.05 | 0.02 | -0.27 | -0.32 | -0.22 | -0.14 | -0.20 | -0.09 |
| Texas | 0.10 | 0.08 | 0.12 | -0.30 | -0.33 | -0.27 | -0.02 | -0.04 | 0.01 |
| Utah | 0.22 | 0.17 | 0.26 | 0.03 | -0.03 | 0.09 | 0.17 | 0.09 | 0.25 |
| Vermont | 0.14 | 0.03 | 0.25 | 0.21 | 0.07 | 0.35 | -0.15 | -0.32 | 0.02 |
| Virginia | 0.09 | 0.06 | 0.13 | 0.12 | 0.08 | 0.15 | -0.02 | -0.06 | 0.02 |
| Washington | 0.07 | 0.04 | 0.10 | 0.02 | -0.03 | 0.07 | -0.06 | -0.10 | -0.01 |
| West Virginia | -0.11 | -0.17 | -0.04 | -0.13 | -0.19 | -0.06 | -0.13 | -0.22 | -0.03 |
| Wisconsin | 0.17 | 0.13 | 0.21 | 0.03 | -0.02 | 0.09 | -0.09 | -0.15 | -0.04 |
| Wyoming | 0.12 | 0.02 | 0.23 | 0.12 | 0.00 | 0.24 | 0.10 | -0.08 | 0.27 |

Source 2015 ACS PUMS
D.C. refers to the District of Columbia
CI refers to the confidence interval

Taking into consideration that the Census Bureau divides the United States into four regions, Northeast, Midwest, South and West, the scores in Table 9.1 show that for two of the three dimensions, a state from the southern region has the lowest average score (West Virginia and the District of Columbia). The states with the highest average score in each dimension come from different regions. Exploring the average scores further by dimension, we see that for 'labour force participation', four of the five states with the lowest average scores are situated in the south (West Virginia, Arkansas, Alabama, and Florida). On the other hand, four of the five states with the highest average scores are in the Midwestern region (North Dakota, South Dakota, Nebraska, and Minnesota). For 'economic security,' three of the five states with the lowest average scores are from the southern region (Mississippi, Louisiana and Texas). At the higher end of the average scores for 'economic security,' three of the five states with the highest average scores are situated in the northeast region (New Hampshire, Massachusetts, and Connecticut). Finally, for the 'marriage as a social resource' dimension, three of the five states with the lowest average score are in the south (Louisiana, Florida, and the District of Columbia). At the higher end, the scores are less geographically concentrated, though two of the five states with higher average scores are situated in the western region (Utah and Wyoming).

Analysing these basic scores, there are two things worth noting. Firstly, though the results show that a southern state has the lowest score in each dimension, not one southern state consistently has the lowest average score across dimensions. Secondly, these scores represent just one year of analysis. Therefore, no generalisations are made regarding the states and its respective score in each dimension of (dis)advantage.

## 9.4 Multilevel results: Determining the existence of state-level variation in (dis)advantage

The primary task of this chapter and the reason for utilising multilevel modelling is to explore if there is variation in disadvantage across the United States. As recognised in section 6.4.1, this information is generated from estimating unconditional (or null) variance components models, which contain no state or individual level predictors. The results for each dimension of (dis)advantage are presented in Table 9.2. The models suggest that 'economic security' has the highest average score across states with a score of 0.030 ± 0.054.[81] The average scores for 'labour force participation' (0.005 ± 0.023) and 'marriage as a social resource' (0±0.025) are quite nearly zero, indicating that across states the averages remain quite close to the average values uncovered in Chapter 7 (Table 7.4).

The unconditional models in Table 9.2 do provide evidence that there are between state differences in each dimension. This is determined by two things as recognised in Chapter 6 (section 6.4.1): the estimated variation in the random components, which represent the state, and the estimations of the degree of non-independence across individuals in each outcome variable (Hayes,

[81] In chapter 8 of this thesis, it was recognised that the products of a factor analysis are continuous measures with a mean of zero. Therefore, it is not surprising that the means for these are quite close to zero.

2006). Firstly, the estimated variance of the random components of the null models are 0.007 ('labour force participation'), 0.039 ('economic security'), and 0.007 ('marriage as a social resource'). Each are statistically different from zero ('labour force participation:' z=3.5, p<.001; 'economic security:' z=4.875, p< .001; 'marriage as a social resource:' z=3.5, p<.001). This indicates that there is evidence of differences between states in each dimension of (dis)advantage. Secondly, the degree of non-independence in each respective outcome variable is tested via the intraclass correlation coefficient (ICC), as acknowledged in the methodology chapter (section 6.4.1). The ICC for 'labour force participation' is 0.007, signifying that 0.7% of the total variance in labour force participation is attributed to differences between states. Accordingly, 2.4% of the variance in 'economic security' and 0.03% of the variance in 'marriage as a social resource' is accounted for by differences between states. From the results of the null model, we can see that more variation (more than twice as much) in 'economic security' can be attributed to differences between states than the other dimensions of (dis)advantage.

**Table 9. 2: Unconditional multilevel results for each factor of (dis)advantage**

| | Labour Force Participation | Economic Security | Marriage as a social resource |
|---|---|---|---|
| **Fixed Effects** | | | |
| Level 1 | | | |
| Intercept | 0.005 (.012) | 0.030 (.028) | 0 (.013) |
| **Random Effects** | | | |
| Level 2 (state) | | | |
| Intercept | 0.007 (.002)*** | 0.039 (.008)*** | 0.007(.002)*** |
| Residuals | 1.052 (.003)*** | 1.606 (.005)*** | 2.014 (.006)*** |
| Model summary | | | |
| -2(LL)[82] | 678,419.66 | 777,986.92 | 831,007.18 |

Source: 2015 ACS PUMS
Standard errors in parentheses
*p<0.05, **p<0.01, *** p<0.001

[82] LL refers to the log likelihood. -2LL represents the deviance addressed in Chapter 6.

I recognise that the ICCs for each of the null models are quite close to 0, below an often cited cut off of 0.05, or 5% (recognised in Glaser and Hastings, 2011). It has been suggested that a multilevel model is not necessary when the ICC is low because it implies that individuals (the level-1 units) are statistically independent and subsequently, single-level models can be effectively used (e.g. Thomas and Heck, 2001). Despite this, I choose not to return to a single level model for several reasons. Firstly, the research questions acknowledged in section 9.1 call for a methodological approach that simultaneously examines multiple levels of analysis in order to explore the variation in disadvantage. There are 51 states at the second level in the models. Should I revert to a single-level model, I would have to include 50 additional parameters (with one left out for reference) to estimate the variances. This can be effectively estimated with one parameter in a multilevel model. Next, there have been instances in which a low value of the ICC is deemed okay, beneficial, and/or worth exploring further (Hayes, 2006; Heinrich and Lynn, 2001; Huang, 2018; Kreft and de Leeuw, 1998). For instance, Kreft and de Leeuw (1998) point out that a small ICC can inflate the level of alpha in a single level analysis; they subsequently advise not to utilise a single-level model even if the ICC is low. They acknowledge that when the number of observations within a group (in this case, state) is large (greater than or equal to 100), a small ICC can inflate the Type I error rate[83] (Kreft and de Leeuw, 1998) in an OLS regression.[84] Given that there are 234,846 sample members in the data set across 51 states, there is an average of over 4,000 individuals per state. Therefore, we can be reasonably assured that in the single-level models, the alpha levels could also be inflated. Thirdly, the use of the OLS regression is discouraged with nested data,[85] because it produces misleading standard errors and incorrect degrees of freedom for clustered coefficients (Huang, 2018). It also violates the assumption of

[83] In their example, Kreft and de Leeuw (1998) find that the Type 1 error rate is inflated from the assumed alpha of 0.05 to 0.17 with n=100.

[84] Huang (2016) acknowledges that with clustering, the standard errors produced from an OLS regression may also be larger, leading to increased Type II errors which results in a failure to reject the null hypothesis.

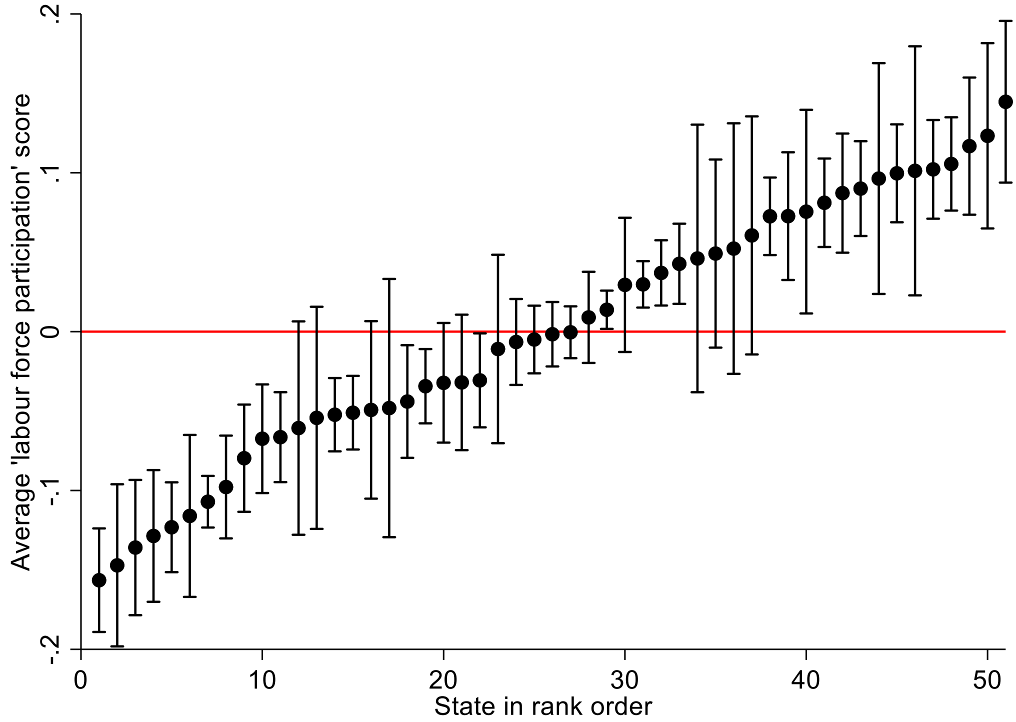
[85] Note that in the previous chapter, the unit of analysis is the individual. An OLS regression analysis was deliberately chosen to answer the research question. For this chapter, the multilevel model addresses this relevant research question and considers simultaneously multiple levels of analysis and their relationship to the respective outcome variables (Pike and Rocconi, 2012).

independence used in OLS regression analysis (Peugh, 2010). The issue of non-independence was addressed in section 6.3. Lastly, other tests and statistics examining the suitability for multilevel modelling of the PUMS data proved statistically significant. These tests are 1) the likelihood ratio test and 2) the design effect. A likelihood ratio test, which determines the contribution of an additional level of analysis by comparing the fit of the models with and without the additional level (Bolker *et al.*, 2009), was conducted for each null model. Each test was statistically significant (labour force participation: $\chi(1)^2 = 877.32$, $p < 0.001$; economic security $\chi(1)^2 = 3659.53$, $p < 0.001$; marriage as a social resource: $\chi(1)^2 = 321.52$, $p < 0.001$), indicating that a multilevel model is preferable over a single level model in each case. In addition, the design effect favours the use of MLM. The design effect statistics are 33.23, 111.5, and 14.82 for 'labour force participation,' 'economic security,' and 'marriage as a social resource,' respectively. Each value is substantially higher than the recommended value of 2 (Peugh, 2010). The results of these tests suggest that a multilevel model is sufficient over a single level model.

As the null models provide evidence of some between state variance in each dimension of (dis)advantage, this can also be produced graphically. Figures 9.1, 9.2, and 9.3 present 'caterpillar plots' of the state effects of the null model in rank order for 'labour force participation,' 'economic security,' and 'marriage as a social resource' respectively. 51 level 2 (state) residuals are plotted for each state in the data set. Recall from 6.4.1 that the residuals represent state departures from the average score in the respective factor of (dis)advantage. Each plot is presented with 95% confidence intervals. So we are able to tell that the majority of states differed significantly from the average line with an alpha of 0.05.

I am making a conscious decision not to name the states in rank order for any of the dimensions in the respective plots. The average scores are listed in Table 9.1. Because I have recognised the importance of place in section 2.5, I instead acknowledge the states' geographic location. I start with 'labour force participation.' For instance, five of the top-scoring states in labour force participation are situated in the Northeast United States. At the other end of the spectrum, eight of the ten states ranked lowest are in the southern parts of the United States.

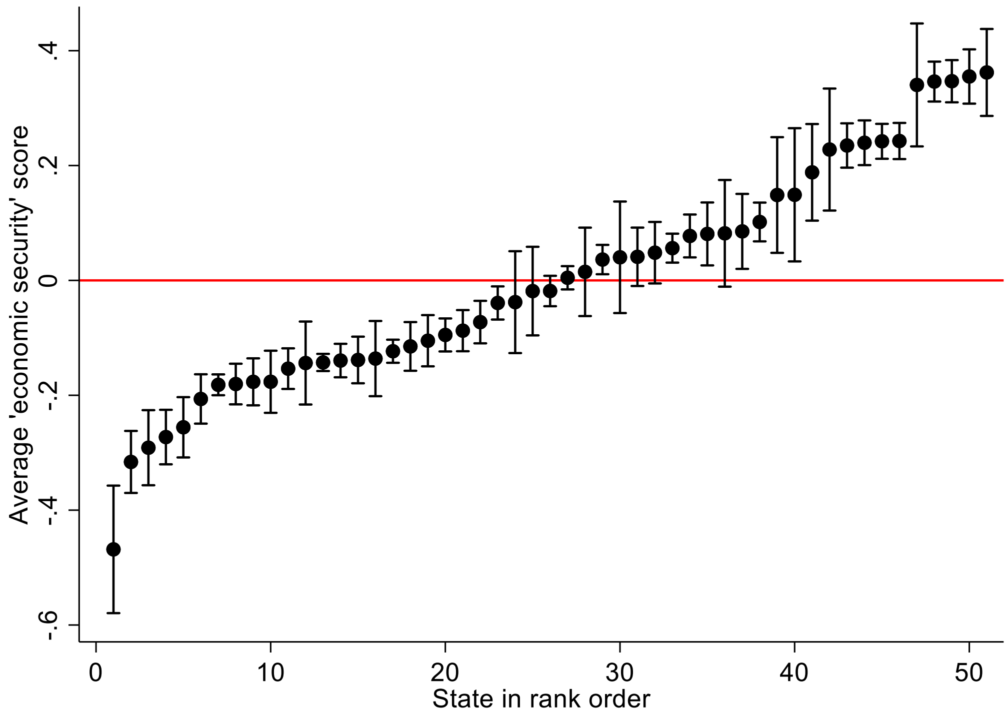
**Figure 9.1:'Caterpillar plot' examining state effects on 'labour force participation'**




Source: 2015 ACS PUMS

Figure 9.2 presents the 'caterpillar plot' of state-level residuals for 'economic security.' The states with the highest scores in 'economic security' are primarily in the Northeast. Those ranking lowest are in the south. The intervals for these ranks, compared to the intervals in Figure 9.1 for 'labour force participation' is much smaller, indicating that for more states in the economic security dimension, we can be 95% confident of their departure from the average score.

**Figure 9.2: 'Caterpillar plot' examining state effects on 'economic security'**



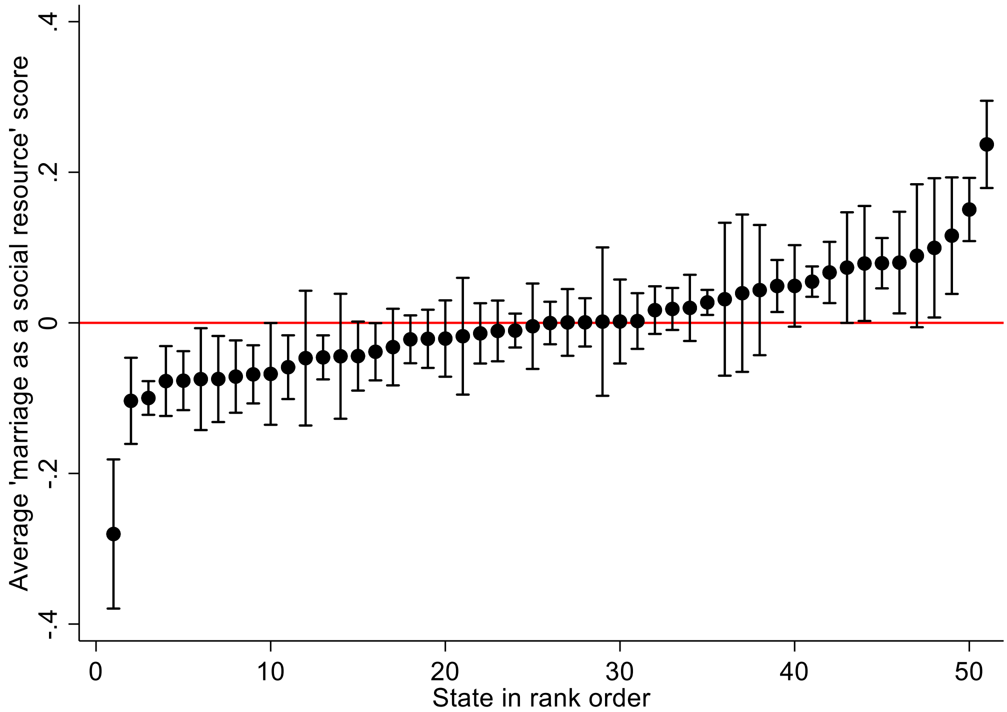

Source: 2015 ACS PUMS

Table 9.3 presents the 'caterpillar plot' for the 'marriage as a social resource' dimension. Here, many more states do not vary significantly from the average. Figure 9.3 does indicate, however, that the top three scores are in different geographic areas of the United States: West, Midwest, and Hawaii (Pacific).[86] Two of the three lower-scoring states are in the southern United States.

These geographic plots provide a visible representation of the differences between states in each dimension of (dis)advantage. The plot for 'economic security' supports the findings in Table 9.2 that there is more variation between states than the other factors. The plot for 'marriage as a social resource' is quite flat, indicating that there is less variation across states, compared to the other two dimensions.

[86] Hawaii and Alaska are interesting cases, because they are not on the continental part of the United States.

**Figure 9. 3: 'Caterpillar plot' examining state effects on 'marriage as a social resource'**




Source: 2015 ACS PUMS

## 9.5 Multilevel results: Controlling for sociodemographic characteristics

The null models provide some generally useful information about each dimension of (dis)advantage. From just those models, we find evidence that variation exists across states for each dimension of (dis)advantage. The first research question of this chapter was successfully answered. Subsequently, the null models are expanded to random intercepts models to include individual-level characteristics as recognised in section 6.4.1b. Here, I can determine if the variation found in the null models hold whilst controlling for these characteristics. As acknowledged in section 9.2, two sets of random intercepts models are examined. The first random intercepts model controls for non-intersectional sociodemographic characteristics, age, race, gender, and citizenship status. The second set of random intercepts models explore intersectionality and control for the intersectional characteristics, age and gender. I will explore the non-intersectional models first.

### 9.5.1 Controlling for non-intersectional sociodemographic characteristics

The results of the expanded models with non-intersectional characteristics are presented in Table 9.3. There are a few noteworthy results. Firstly, across all three models, the coefficients [87] and significance levels for the independent variables are fairly consistent with the OLS regression coefficients presented in Chapter 8 (Table 8.3). For instance, women have less advantage in each category than men, on average. To see the similarities in coefficients for the OLS and MLM models is common, as Huang (2018) acknowledges. It could be inferred from this that the MLM is not required but it is necessary because I utilise it to answer a specific research question about variance across states that cannot be answered via an OLS regression.

With the addition of these variables, the between state and within state (that is, between individual) variance in each dimension of (dis)advantage decreased, though not by much. The most substantial change was in the 'labour force participation' model. The ICC decreases to .005. This signifies that after controlling for the sociodemographic variables, 0.5% of the variation in 'labour force participation' is explained by differences between states. This is 0.2% lower than the null model. Because the value dropped so little, it suggests that a lot of the variance in this dimension has a lot to do with individuals. There is, however, still some variation that is unique to the state. Interestingly, the value of the ICC increases in the 'economic security' model to 0.027 (an increase of 0.003 from the null model). This result suggests that after controlling for the sociodemographic variables, 2.7% of the variance in 'economic security' is between states. The ICC remains unchanged for the 'marriage as a social resource' model, indicating that after controlling for race, gender, age, and citizenship status, 0.03% of the variation in the 'marriage as a social resource' dimension is attributed to the

[87] Recall from the methodology chapter that none of the predictor variables are centred. This follows the advice of Nezlek (2012) who recommends that categorical variables can be entered into multilevel models un-centred.

difference between states. Though small, the results indicate that there is still variation in (dis)advantage across the United States, beyond the compositional effects of the individual sociodemographic variables.

**Table 9. 3: Multilevel analysis of the non-intersectional sociodemographic characteristics on the dimensions of (dis)advantage**

| | Labour Force Participation | Economic Security | Marriage as a Social Resource |
|---|---|---|---|
| **Fixed Effects** | | | |
| Level 1 | | | |
| Intercept | -0.789 (.01)*** | 0.318 (.028)*** | 0.134(.013)*** |
| ***Age*** | | | |
| 65 and up (ref) | | | |
| 55-64 | 0.98 (.006)*** | 0.15 (.008)*** | 0.03 (.009)** |
| 45-54 | 1.33 (.006)*** | 0.16 (.008)*** | 0.03 (.009)** |
| 35-44 | 1.37(.006) *** | 0.03 (.008) ** | 0.08 (.01)*** |
| 25-34 | 1.36(.006)*** | -.25 (.008) *** | -0.17 (.01)*** |
| 18-24 | 1.04(.007) *** | -.77 (.009)*** | -0.56 (.011)*** |
| ***Race*** | | | |
| White alone (ref) | | | |
| Black alone | -0.10 (.006) *** | -0.60 (.009) *** | -0.43 (.01)*** |
| Native | -0.27 (.018)*** | -0.98 (.025)*** | -0.20 (.029)*** |
| Asian alone | -0.05 (.009) *** | 0.40 (.013)*** | 0.13 (.015)*** |
| Pacific Islander | -0.03 (.047) | -0.28 (.063)*** | -0.11 (.075) |
| Other race alone | -0.01 (.011) | -0.76 (.014)*** | -0.04(.017)* |
| Mixed race | -0.08 (.013) *** | -0.26 (.018)*** | -0.24(.021)*** |
| ***Sex*** | | | |
| Male (ref) | | | |
| Female | -0.21 (.004)*** | -.18 (.005)*** | -.13(.006)*** |
| ***Citizenship Status*** | | | |
| Born in the U.S. (ref) | | | |
| Born in U.S. territories | -0.17 (.025) *** | -0.73 (.034)*** | -0.14 (.04)** |
| Born abroad of U.S. parents | 0.05 (.019) ** | -0.01(.026) | 0.07(.031)* |
| Naturalised | 0.08 (.008) *** | -0.37(.01)*** | 0.28 (.012)*** |
| Not a U.S. citizen | -0.11 (.008) *** | -1.14 (.011)*** | 0.42 (.013)*** |
| **Random Effects** | | | |
| Level 2 (state) | | | |
| Intercept | .004 (.001 )*** | .038 (.008)*** | .006(.002)** |
| Residuals | .751 (.002)*** | 1.373 (.004)*** | 1.939(.006)*** |
| Model summary | | | |
| -2(LL) | 599,195 | 741,058.66 | 822,116.14 |

Source: 2015 ACS PUMS
N=234,846
Standard errors in parentheses
*p<0.05, **p<0.01, *** p<.001

The results in Table 9.3 show significant variation across states in each dimension of disadvantage ('labour force participation' and 'economic security': *p<0.001*; 'marriage as a social resource:' *p<0.003*). Examining between state variance, I find that most of the variance between states, while controlling age, gender, race, and citizenship status, is higher in the 'economic security' dimension (0.038 compared to 0.004 in 'labour force participation' and 0.006 in 'marriage as a social resource.' The level 1 variance for each of the models have also decreased, but still statistically significant, *p<0.001*), with the largest decrease occurring in the 'labour force participation' model (1.052 to 0.751). This indicates that some of the within-state variance observed for each dimension of (dis)advantage in Table 9.1 within states may have been due to the variation in sociodemographic characteristics at the individual level.

In order to determine how much variation in the dimensions of (dis)advantage is accounted for by the sociodemographic characteristics, the 'variance accounted for' measure is calculated (acknowledged section 6.4.2). With the addition of these variables to the null model, the residual variance is changed by a factor of 0.714.[88] This produces a 'variance accounted for' measure of 0.286 (equalled to 1 – 0.714). Essentially, 28.6% of the variance in 'labour force participation' is explained by the sociodemographic variables after accounting for the differences between states. This is nearly 3% more of the variation in 'labour force participation' that was not accounted for in the OLS regression presented in Table 8.3. The results indicate that characteristics beyond an individual's control explain nearly one-third of the variance in participating in the labour market. For 'economic security,' the residual variance is changed by a factor of 0.855, indicating a 'variance accounted for' measure of 0.145. The sociodemographic variables explain 14.5% of the variance unexplained by differences between states for the 'economic security' dimension. This is nearly 2% lower than the variance accounted for by the OLS regression in the last chapter. Lastly, for the 'marriage as a social resource' dimension, the residual variance is changed by a factor of 0.963, producing a 'variance

[88] This value is the quotient of the random intercepts model's within state variance and the null model's within state variance.

accounted for' measure of 0.037. Interestingly, compared to the results of the OLS regression presented in the previous chapter, the variance explained is nearly identical (3.8% in the OLS model). The comparatively small variance explained by the sociodemographic variables for this dimension, 3.7%, after accounting for state-level differences, suggests that there are other predictors, unaccounted for by the variables discussed here. These results provide further evidence of not neglecting the multilevel nature of (dis)advantage due to the low values of the ICC.

In order to be sure the variables added to the model does indeed improve the fit of each multilevel model, the deviance statistic is calculated.[89] For each of the dimensions, there is a 16 parameter difference[90] between the unconditional and conditional models. Subsequently, each model is distributed as chi-square with 16 degrees of freedom. Each deviance statistic is statistically significant: labour force participation $\chi^2(16) = 79{,}224.66$, $p<0.001$; 'economic security' $\chi^2(16) = 36{,}928.26$, $p<0.001$; 'marriage as a social resource' $\chi^2(16) = 8{,}891.04$, $p<0.001$. This indicates that with the addition of these individual sociodemographic characteristics, each multilevel model is improved.

In this section, I found evidence that variation between states exists when controlling for non-intersectional sociodemographic characteristics. In the next section, I determine if the variation between states exists whilst controlling for intersectional characteristics.

### 9.5.2 Controlling for intersectional sociodemographic characteristics

Table 9.4 presents the results of the three multilevel models that expand the null models presented in Table 9.1. I include the variables representing race, gender, and their interactions. The likelihood ratio tests, which assess the advantage of a multilevel model over a single-level model, were conducted and were all statistically significant ('labour force participation:' $\chi(1)^2 = 784.55$, $p < 0.001$; 'economic security:' $\chi(1)^2 = 3237.8$, $p < 0.001$; 'marriage as a social resource:' $\chi(1)^2 = 309.69$, $p < 0.001$).

[89] As acknowledged in the methodology chapter, the deviance statistic is equalled to the deviance of the null model minus the deviance of the bigger model (Hayes, 2006). Deviance is equalled to -2 multiplied by the log likelihood.
[90] 16 variables were added to the conditional model representing age, sex, race, and citizenship status.

This indicates that a multilevel model with these variables at level 1 is preferable over a single level model.

**Table 9.4: Multilevel analysis of the intersectional effect of race and sex on dimensions of (dis)advantage**

| | Labour Force Participation | Economic Security | Marriage as a Social Resource |
|---|---|---|---|
| Fixed Effects | | | |
| Level 1 | | | |
| Intercept | 0.134 (.012)*** | 0.239(.026)*** | 0.09 (.012) *** |
| ***Race*** | | | |
| White (ref) | | | |
| Black | -0.13(.011)*** | -0.698(.013)*** | -0.34 (.015) *** |
| Native | -0.23(.031)*** | -1.072(.038)*** | -0.22 (.043)*** |
| Asian | 0.12 (.014)*** | -0.085 (.017)*** | 0.31 (.020) *** |
| Pacific Islander | 0.21(.079)** | -.55(.096)*** | 0.23 (.109) * |
| Other race | 0.26(.017)*** | -1.303(.021)*** | 0.06 (.024) * |
| Mixed race | 0.07(.022)** | -0.475(.026)*** | -0.34 (.030) *** |
| **Sex** | | | |
| Male (ref) | | | |
| Female | -0.26(.005)*** | -0.183(.006)*** | -0.11 (.007)*** |
| ***Interactions (race and sex)*** | | | |
| White women (ref) | | | |
| Black women | 0.20(.015)*** | 0.107(.018)*** | -0.17 (.020)*** |
| Native women | 0.15(.042)*** | 0.115(.051)* | -0.22 (.058) |
| Asian women | -0.05(.019)** | 0.07(.023)** | 0.06 (.03) * |
| Pacific Islander women | -0.12(.108) | -0.118(.131) | 0.19 (.15) |
| Other race women | -0.16(.024)*** | 0.10 (.028)*** | 0.02(.032) |
| Mixed Race women | 0.06(.030)* | 0.07(.037)* | 0.07(.042) |
| Random Effects | | | |
| Level 2 (state) | | | |
| Intercept | 0.006 (.001)*** | 0.033(.007)*** | 0.005 (.001)*** |
| Residuals | 1.034(.003)*** | 1.511(.004)*** | 1.985 (.006)*** |
| Model summary | | | |
| -2(LL) | 674368.96 | 763670.54 | 827533.36 |

Source: 2015 American Community Survey Public Use Microdata Files
N=234,846
Standard errors in parentheses
*p<0.05, **p<0.01, *** p<.001

The results in Table 9.4 show statistically significant variation between states in each dimension (*p<.001*). Looking at the between state variance for 'labour force participation,' we see that it is higher in the intersectional model compared to the model with non-intersectional characteristics. It is lower in comparison to the null model, which is to be expected as we are controlling for characteristics at the individual level. The increase in variance compared to the non-intersectional model suggests that for the variables included in Table 9.4, there are more differences between states in this dimension. For the 'economic security' and 'marriage as a social resource' models, the between state variances in the intersectional models are lower in both the non-intersectional and null models. The within state variances for all models decreased compared to the null models. This suggests that for these characteristics included in Table 9.4, there is less state variation in these models.

Using the 'variance accounted for' measure, I can test again how much of the variance in (dis)advantage is accounted for by the intersectional characteristics and their components. The 'variance accounted for' measure for 'labour force participation' is 0.017, indicating that 1.7% of the variance in this dimension is accounted for race, gender, and its interactions. For the 'economic security' model, 5.9% of the variance is accounted for by intersectional effects of race and gender, after removing between state variance in economic security. For the 'marriage as a social resource' model, 1.4% of the variance in 'marriage as a social resource' is accounted for by the intersectional variables.

The 'variance accounted for' measures for the models in Table 9.4 uncovered some exciting findings. Firstly, compared to the non-intersectional model, the individual characteristics in the intersectional model account for over 20% less of the variation in the 'labour force participation' model. This finding suggests two things. More variation in 'labour force participation' has to do with the other individual-level characteristics that are not included in Table 9.4, including age and citizenship status. Also, this finding does not mean that individuals with the characteristics highlighted in Table 9.4 are more advantaged. For instance, it implies that the average (dis)advantage that women have in this factor compared to men is more consistent across states. Secondly, I find that for these sociodemographic variables, more variance is explained in 'economic security.' In the non-intersectional models, more variance is explained in 'labour force participation.' This is similar to the findings uncovered in the OLS regression, the results which are presented in Table 8.5. For the intersectional OLS models, more variance was explained in 'economic security,' highlighting that these characteristics (race, gender, and the intersection between the two) explain more variation in this dimension of (dis)advantage than it does in the others.

## 9.6 The effect of gender in explaining the variation in (dis)advantage

The models presented in Tables 9.3 and 9.4 offer evidence that there is some variation across states in each dimension of (dis)advantage while controlling for various sociodemographic characteristics. However, these models assume that the slopes for each of these characteristics are fixed among states. There is an assumption that the relationship between an independent variable and the outcome is the same across states. For instance, it assumes that across states, on average, women score 0.26 less than men in the 'labour force participation' dimension. However, the effect of any of the individual characteristics may vary from state to state. Employing random coefficient models, as discussed in section 6.4.1c, I am able to investigate if the relationship between these individual characteristics and (dis)advantage vary across states.

Ideally, the random coefficient models would include each of the individual sociodemographic variables as random, not fixed effects. However, Peugh (2010) cautions that doing this can lead to decreased statistical power and errors in parameter estimation. When attempted for this thesis, a parameter estimation error did occur in the form of non-convergence for both the non-intersectional and intersectional random coefficient models.

The random effect structure is normally distributed with a mean 0 and has a variance-covariance component (Eager and Roy, 2017). Convergence is a result of the iterative algorithm used to solve for the variance-covariance and the coefficients (β's in Equations 6.7). Eager and Roy (2017) acknowledge that convergence consists of many tests that involve checking that the variance-covariance is positive-semidefinite, meaning that no linear transformation of the random effects has a negative variance estimate (an invalid statistical estimate). Peugh (2010) recognises that non-convergence occurs if the maximum number of parameter (β) estimation iterations is reached but a non-negligible change in the value of the log-likelihood between the final two estimation iterations occur. He offers a few reasons for non-convergence: 1) small sample size, 2) imbalanced data, meaning that there are small numbers of individuals from some states, and 3) model misspecification, meaning that a model is being estimated with excessive numbers of level 1 random effects (Peugh, 2010). The latter cause seems to be the cause here for non-convergence, particularly as a random coefficient model would include sixteen additional random effects in the non-intersectional model and fifteen in the intersectional model. This can be remedied by decreasing the number of model's level 1 random effect estimates (Peugh, 2010). A decision is made not to lump the many categories of the various sociodemographic variables. This would result in a loss of information that has already been found in the differences in each of the various categories. Instead of focusing on each of these sociodemographic characteristics, I focus on gender. [91] Compared to men, women have a lower score in each model on average. Additionally, it is a sociodemographic variable found in both the intersectional and non-intersectional random intercept models. The results for the random coefficient models that contain a random effect for gender are presented in Table 9.5.

[91] Race was also considered for the random coefficient model and tried in Stata 15. After more than 24 hours continuously running, the model failed to converge. This is likely because there are six categories representing race. Though Peugh (2010) acknowledges simplifying the model to remedy non-convergence, this is not ideal for this variable. Based on the results in this and the previous chapter, we find that there are differences between the categories and there would be loss of information in lumping all minority groups into one to compare to the reference group, White.

**Table 9. 5: Random coefficient multilevel models with 'gender' as a random effect**

| | Labour Force Participation | Economic Security | Marriage as a Social Resource |
|---|---|---|---|
| Fixed Effects | | | |
| Level 1 | | | |
| Intercept | -0.793 (.01)*** | 0.317 (.029)*** | 0.129 (.013)*** |
| | | | |
| ***Age*** | | | |
| **65 and up (reference)** | | | |
| 55-64 | 0.98 (.006)*** | 0.15 (.008)*** | 0.03 (.009)** |
| 45-54 | 1.33 (.006)*** | 0.16 (.008)*** | 0.03 (.009)** |
| 35-44 | 1.37(.006) *** | 0.03 (.008) ** | 0.08 (.01)*** |
| 25-34 | 1.36(.006)*** | -0.25 (.008) *** | -0.17 (.01)*** |
| 18-24 | 1.04(.007) *** | -0.77 (.009)*** | -0.56 (.011)*** |
| | | | |
| ***Race*** | | | |
| **White (reference)** | | | |
| Black | -0.10 (.006) *** | -0.60 (.009) *** | -0.43 (.01)*** |
| Native | -0.27 (.018)*** | -0.98 (.025)*** | -0.20 (.029)*** |
| Asian | -0.05 (.009) *** | 0.40 (.013)*** | 0.13 (.015)*** |
| Pacific Islander | -0.03 (.047) | -0.28 (.063)*** | -0.11 (.075) |
| Other race | -0.01 (.011) | -0.76 (.014)*** | -0.04(.017)* |
| Mixed race | -0.08 (.013) *** | -0.26 (.018)*** | -0.24(.021)*** |
| | | | |
| ***Sex*** | | | |
| ***Male (reference)*** | | | |
| Female | -0.20 (.007)*** | -0.17 (.007)*** | -0.12(.009)*** |
| | | | |
| ***Citizenship Status*** | | | |
| **Born in the U.S. (reference)** | | | |
| Born in U.S. territories | -0.17 (.025) *** | -0.73 (.034)*** | -0.14 (.04)** |
| Born abroad of U.S. parents | 0.05 (.019) ** | -0.01(.026) | 0.07(.031)* |
| Naturalised | 0.08 (.008) *** | -0.37(.01)*** | 0.28 (.012)*** |
| Not a U.S. citizen | -0.11 (.008) *** | -1.14 (.011)*** | 0.43 (.013)*** |
| | | | |
| Random Effects | | | |
| Level 2 (state) | | | |
| Intercept | 0.004 (.001 )*** | 0.042 (.009)*** | 0.005(.001)*** |
| Residuals | 0.750(.002)*** | 1.372 (.004)*** | 1.939(.006)*** |
| Coefficient | 0.002(.001)* | 0.001(.001) | 0.002(.001)* |
| Covariance | 2.13e-06 (.001) | -.004(.002)* | 0.001(.001) |
| ***Model summary*** | | | |
| -2(LL) | 599,110.16 | 741,046.02 | 822,104.66 |

Source: 2015 ACS PUMS
N=234,846
Standard errors in parentheses
*p<0.05, **p<0.01, *** p<.001

The similarity in the results (fixed and random components) for 'labour force participation' in Table 9.5 and Table 9.3 suggests that there is no gender effect in this dimension across states. However, the random intercept components are statistically significant ($p<0.001$). This indicates significant between state variation in 'labour force participation,' with respect to the average difference between men and women. In addition, the variance across states in the effect of being female is also statistically different from zero ($p<0.05$). So it would seem that the effect of being female in the 'labour force participation' dimension varies across states. The value of the covariance is positive, quite close to zero, but also statistically insignificant, ($p>0.05$). However, as recognised in Chapter 6, section 6.4.1c, I employ a likelihood ratio test as it is noted to be a more accurate test of the random effects (Hayes, 2006; Peugh, 2010). As the random intercept model in Table 9.3 is nested in the random coefficient model in Table 9.5, the deviance of the random coefficient is subtracted from the deviance of the random intercepts model. The test statistic is then 84.84 at 2 degrees of freedom because there is a two-parameter difference. The likelihood ratio test $\chi(2) = 84.84$, $p<.001$ is statistically significant, indicating that the relationship between gender and 'labour force participation' is not the same for each state. It also provides evidence that for this sociodemographic characteristic, a random coefficient model is preferred over a random intercepts model.

For economic security, the regression coefficients in the fixed part of the model are similar to the results for 'economic security' in Table 9.3. The random intercept components are statistically significant ($p<0.001$). This indicates significant between state variation in 'economic security,' with respect to the average difference between men and women. Interestingly, the variance across states in the effect of being female is statistically insignificant. The covariance for this dimension is negative and statistically significant ($p<0.05$). The negative covariance suggests that states with lower scores in female 'economic security' would have a higher difference between men and women in this dimension. The value of the random effects correlation is -0.618. This indicates that the coefficient for 'female' (which is the average difference between men and women) is larger among states with a lower average score in this dimension. Essentially, a state ranking lower than the average is likely to have large differences in 'economic security' between men and women. The likelihood ratio test, $\chi(2) = 12.64$, $p<0.05$, is statistically significant. This signifies that with a 95% confidence that the relationship between gender and 'economic security' varies across states.

For the final model in Table 9.5, the results for 'marriage as a social resource' showed identical coefficients as presented for the respective model in Table 9.3. The random intercept components are statistically significant ($p<0.001$), just as they were in the other two models. Similar to the 'labour force participation' model, the variance across states in the effect of being female is statistically different from zero ($p<0.05$) and the value of the covariance is positive, quite close to zero, and statistically insignificant, ($p>0.05$). The likelihood ratio test, however, is statistically significant ($\chi(2) = 11.48$, $p<0.05$). This signifies that like 'economic security' and 'labour force participation,' the relationship between gender and the 'marriage as a social resource' dimension varies significantly across states.

This section has provided the answer to the last research question posed in this chapter: **Does the relationship between individual sociodemographic characteristics and multidimensional disadvantage vary significantly across the United States?** The results of the random coefficient models offer evidence that for gender, there is a significant variation across states. For each dimension of (dis)advantage, the effect of gender depends on the states in which that individual resides. For the 'labour force participation' dimension, there is a 99% confidence in this relationship. For both the 'economic security' and 'marriage as a social resource' dimension, there is a 95% confidence of the variation in the relationship between gender and these dimensions across states.

## 9.7 Conclusion

Utilising multilevel models, this chapter delved into exploring three things: 1) if there was significant variation across states in the multiple dimension of disadvantage 2) if that variation continued taking into consideration individual sociodemographic characteristics, and 3) if the relationship between sociodemographic characteristics and the dimensions vary across states. The results of twelve multilevel models offer significant evidence that there is variation across states in each dimension of disadvantage, even controlling for intersectional and non-intersectional sociodemographic characteristics. The random intercepts models (presented in Table 9.3) indicated that nearly one-third of the variation in 'labour force participation' and nearly 15% in 'economic security' is accounted for by the sociodemographic characteristics, age, race, gender, and citizenship status, after accounting for the differences between states. If we consider that the 'labour force participation' dimension consists of actual participation in the labour market and ability to participate, it is an interesting finding that implies that some of the disadvantages faced by those least advantaged, women and minorities, are structural in nature and consistent across states. The evidence that 'economic security' composes some structural disadvantages is further highlighted in the intersectional model (Table 9.4). After accounting for the variation across states, more variation in 'economic security' (6%) is explained by the intersectional characteristics than in the other dimensions (just under 2% each). The random coefficient models in Table 9.5 indicate that the effect of gender on multidimensional (dis)advantage varies significantly across states. This means that for each dimension of (dis)advantage the effect of being female and the experience of disadvantage differs based on where you live.  This analysis offers concrete evidence that 1) disadvantage is not unidimensional and 2) where you live, particularly for women, matters.

# Chapter 10

# Conclusions

## 10.1 Introduction

In 2015, the United States counted 43 million Americans as income poor (Proctor *et al*., 2016). The measure used to determine this status does not take into account the various manifestations of disadvantage that all Americans may face. Only income is considered in determining poverty status. This thesis reconsidered disadvantage in the United States and sought to quantify multiple forms of disadvantage. By conceptualising disadvantage as social exclusion in the United States, I have set out to understand the various manifestations of disadvantage in the United States that may extend beyond low levels of income. Subsequently, I have empirically uncovered salient dimensions of disadvantage utilising 'big' American data. Furthermore, I have explored how these dimensions relate to individual sociodemographic characteristics, including the intersection between race and gender, and analysed how these relationships might vary across the country.

This research has subsequently added to knowledge by providing insight into the experiences of disadvantage at the individual level. In this chapter, I reflect on my work on the thesis. I divide the chapter into seven sections. Sections 10.2 to 10.4 summarise each of the three findings chapters. Section 10.5 discusses the policy implications of this thesis. Section 10.6 discusses the research limitations. Section 10.7 makes recommendations for further research. Section 10.8 concludes the thesis.

## 10.2 What were the factors of social exclusion in the United States?

The results of the factor analysis discussed in Chapter 7 unveiled three factors of social exclusion: 1) labour force participation, 2) economic security, and 3) marriage as a social resource. It was in this chapter that I was able to test the application of the Bristol Social Exclusion Matrix (B-SEM) (discussed in Chapter 3), a framework of social exclusion developed by Levitas *et al*. (2007). I applied the framework to data from a context in which the concept of social exclusion is comparatively less used. While the B-SEM had been used in papers for the United Kingdom Cabinet Office (Cusworth *et al.*, 2009; Oroyemi *et al*., 2010), to the best of my knowledge, it had not been applied to American data prior to this research project. The advantage of applying the framework to American data was to test if the domains of social exclusion held outside of British and other European contexts.

The B-SEM identifies three interconnected domains: 1) resources, 2) participation, and 3) quality of life. Interestingly, only one dimension of (dis)advantage in my analysis, 'marriage as a social resource' aligned with the B-SEM. Every indicator for the social resources subdomain of the B-SEM loads onto 'marriage as a social resource.' The other two dimensions, 'labour force participation' and 'economic security' comprised of many different components of the B-SEM. Firstly, 'labour force participation' included indicators from the 'economic participation,' 'health and well-being,' and 'transportation' subdomains, which encompass the participation and quality of life domains of the B-SEM. The 'economic security' dimension of (dis)advantage included indicators from the 'economic resources,' 'cultural capital and participation,' 'living environment and standard of living,' and 'transportation' subdomains. Interestingly, the 'economic security' dimension of (dis)advantage includes indicators from every domain from the B-SEM, which highlights two things. It lends weight to a warning by Lister (2004) not to downplay the importance of income. In addition, this highlights that economic security encompasses more issues than a low level of income. It surrounds issues of health care, overcrowded housing, and food insecurity.

Therefore, I found that the B-SEM framework was helpful in understanding and providing a starting point for measuring social exclusion in the United States. By conducting an empirically driven analysis via the use of exploratory factor analysis, I was able to show that these indicators are measuring dimensions of (dis)advantage in the United States that do not all fall neatly into the resources, participation, and quality of life domains of the B-SEM. Interestingly, Bailey *et al*. (2018), who use the same methodological technique (exploratory factor analysis) on UK data based on the B-SEM found five factors of social exclusion, but also found that their domains did cross groupings. For instance, their first factor, 'economic resources and housing' contained indicators from all B-SEM domains (Bailey *et al.*, 2018). In the following subsections, I summarise the substantial interpretation of the factors derived from my analysis.

In the following subsections, I summarise the substantial interpretation of these factors.

### 10.2.1 Labour force participation

The results of the analysis (Table 7.3) suggests that there is a two-fold understanding of this factor that encompasses actual participation in the labour market and the ability to participate in the labour market. The first understanding embodies the traditional and basic understanding of labour force participation. That is, whether or not an individual is actually in the labour market.

The second component reflects the changing structure of the labour market and of work. It focuses on the ability to participate. This component is recognised by examining the items that loaded highly onto this factor: disability status, driving as a means to transportation to work, and travel time to work. Firstly, having a disability has a negative relationship with this factor. Stapleton *et al.* (2006) offered a reason for this, acknowledging that people with disabilities are often not able to work as they are trapped in policies that provide benefits on the condition of not being able to work. This component of 'labour force participation' also recognised travel time and means of transportation to work. It suggests that the longer it takes an individual to commute to work, the more likely they are to participate in the labour market. Driving to work also has a highly positive relationship with this factor.

The second component of the 'labour force participation' factor also highlights that space matters for labour market outcomes (Fernandez and Su, 2004). The emphasis is on space as a constraining factor for various individuals within the labour market, particularly for racial minorities (Massey and Denton, 1993) as racial segregation has made it difficult for minorities to overcome the mismatch between job location and their homes. It is definitely not a new connection as Holzer (1991) acknowledges that the distance between a central city resident and likely work locations have been increasing over time. It signifies that these issues continue to be an issue of disadvantage in the United States.

This dimension of (dis)advantage suggests that considering solely the rate of employment in the United States is not a reliable indicator to encompass all the reasons an individual might be excluded from full participation in the labour market. It is evident that we have to learn about participation behaviour and recognise the various constraints on employment. Taking two components of labour force participation and analysing them concurrently, this factor offers a more complete picture of (dis)advantage in labour force participation in the United States.

### 10.2.2 Economic Security

Like the 'labour force participation' dimension, I interpret 'economic security' with a two-fold understanding. Interestingly, the variables that load onto this factor embodied a United Nations (2009) definition of economic insecurity that suggests *"economic insecurity rises from the exposure of individuals, communities, and countries to adverse events and from their inability to cope with and recover from the costly consequences of those events"* (quoted in Bossert and D'ambrosio, 2013, p. 1018).

One on hand, three of the six items loading on this factor 1) 'living in overcrowded housing,' 2) 'income poverty,' and 3) being 'food stamps recipient,' reflect the acknowledgement that economic insecurity arises when individuals are exposed to adverse events. These sorts of events put individuals at risk of social exclusion in the form of economic insecurity. On the other hand, the United Nations (2009) acknowledge that economic insecurity arises when individuals are unable to deal with and recover from these adverse events. The remaining three factors further encompass this understanding of economic insecurity.

This dimension of (dis)advantage is particularly interesting because much work has been conducted on economic insecurity (Bossert and D'ambrosio, 2013; Burns and Gimpel, 2000; Catalano, 1991; Dominitz and Manski, 1996; Hacker, 2011; Jacobs, 2007; Mughan, 2007; Osberg, 2015; Osberg and Sharpe, 2014; Rejda and Haley, 2004; Wroe, 2016). Economic security is of interest globally because there is concern over its causes. Mughan (2007), for instance, ties economic insecurity to job insecurity. This concentration somewhat contradicts the findings of my analysis. The respective variable for employment did not load on the 'economic security' dimension. It loaded on the 'labour force participation' dimension. However, the results presented in Table 7.5 recognised a small, positive correlation between 'labour force participation' and 'economic security' (0.22). This indicates that though they are inter-related, there should be some shift toward separating employment from economic security in academic and policy discussions. These dimensions represent distinct issues of disadvantage in the United States.

### 10.2.3 Marriage as a social resource

This factor proved most difficult in interpreting because of the relationships the variables have with it. As shown in Table 7.3, 'marital status: married' had a positive association and 'marital status: divorced' had an equally negative relationship. It is not my intention with this factor to suggest that marriage will lift anyone out of poverty, income or otherwise. Indeed, studies have already shown that if both in the union are income poor, then the benefit of marriage is actually minimal. Sigle-Rushton and McLanahan (2002), for one, argue that marriage proponents are overstating its benefits when they compare the earnings or income poverty rates of single-mother families to those of married, two-parent families.

Instead of relying solely on an economic interpretation of this factor, I acknowledge that marriage and partnerships can provide advantages in the form of social resources. Each of the variables that comprise this dimension of (dis)advantage were suitable indicators for the social resources subdomain of the B-SEM. Therefore, it is appropriate that I interpret this factor as such. Having adequate social capital and resources produces social support and social leverages at the individual level (Johnson *et al.*, 2011). In partnerships, such support and leverage can take the form of access to health care and an expansion of social support networks (Grzywacz and Fuqua, 2000). In contrast, we would expect to see some of the associated benefits of partnerships to be less prevalent amongst those who are divorced. Indeed, the 'marital status: divorced' variable in Table 7.3 had a negative association with this dimension. In addition, research suggests cohabitation arrangements (considered under the 'unmarried partners' variable) are relatively new, brief, and not as financially or socially intertwined, compared to a union by marriage (Garrison, 2005). This suggests that the positive social externalities associated with marriage are less available in unmarried partnerships.[92] Again, the analysis in Table 7.3 supports this finding, elucidating a negative relationship between this factor and 'unmarried partnership.' However, the relationship between 'marriage as a social resource' and 'marital status: divorced' has a stronger negative relationship, compared to the relationship between this factor and 'unmarried partnership.' This offers some evidence that there are some added benefits in the United States with being married that the other types of partnerships are unable to enjoy.

An important point to consider is that some groups of the population are excluded from having the choice of acquiring this particular social resource and the associated benefits. For instance,

[92] It should be noted here again that this analysis only explores one year of data. Therefore, I am not able to comment much on the brevity or longevity of any of the relationships highlighted.

the analysis in Chapter 8 highlighted that Black women often face racial prejudice that keeps them excluded from a marriage with potential spouses who are not Black (Johnson and Loscocco, 2015). Additionally, same-sex couples are still prohibited from marrying in many American states and their partnerships are often not recognised in many legal spaces. Lastly, the income-poor themselves are oftentimes excluded from marrying because of the associated costs of officially marrying. This often leaves many left in a position of cohabiting, facing increased possibilities of lower levels of social resources.

In sum, there are social benefits associated with marriage. It is essential to consider these along with the economic. However, it is recognised that for some groups of the American population, the ability to actually secure these benefits are structurally out of reach.

## 10.3 To what extent were sociodemographic characteristics associated with multidimensional disadvantage in the United States?

In chapter 8, I explored the relationships between individual sociodemographic characteristics and the dimensions of (dis)advantage. This analysis uncovered exciting findings. Firstly, I found that on average, women were less advantaged in each dimension of (dis)advantage than men, holding the other sociodemographic characteristics constant. Secondly, the results suggest that Black, Native, and Mixed-race individuals were less advantaged in every dimension than Whites, on average. Interestingly, compared to Whites, every minority group was less advantaged in the 'economic security' dimension, on average and holding the other variables in the model constant. Next, considering age, the results suggest that while individuals between the ages of 18 and 34 had more advantage in the 'labour force participation' dimension of (dis)advantage, on average, compared to individuals aged 65 and up, they on average had less advantage in the 'economic security' and ‘marriage as a social resource’ dimensions. Interestingly, this follows the same pattern for Black women, where Black women are less advantaged in ‘economic security’ and ‘marriage as a social resource’ than the respective reference group. Finally, focusing on citizenship status, the results suggest that individuals born in American territories were less advantaged, on average than individuals born in the United States. In the ‘economic security’ dimension, naturalised citizens and non-American citizens were also less advantaged than individuals born in the United States. The results provided some insights into the experience of disadvantage at the individual level. It further highlighted that age, gender, race, and citizenship status explained 26% of the variation in ‘labour force participation,’ 16% in ‘economic security, and just under 4% in ‘marriage as a social resource.’ Though I gathered substantial insights from this analysis, I, however, recognised that a focus on singular categories was incomplete.

Instead of relying solely on the singular categories of race, gender, citizenship status and age to explore the individual experience of disadvantage in the United States, I have incorporated intersectionality into my work. Intersectionality recognises that an individual at intersecting identities faces disadvantage not well understood by just looking at each category alone (Bowleg, 2012; Crenshaw, 1991; McCall, 2005). In this thesis, I focused on the intersection between gender and race for two previously stated reasons. Firstly, it was the intersection of characteristics that Crenshaw (1991) used in her articulation of intersectionality. Her research heavily focused on the experience of Black women. Secondly, I recognise the importance of reflexivity in intersectionality research (Johnson and Christensen, 2013). As a minority woman, I sit at the intersection of those characteristics. I am seeking to contribute to a literature that has neglected to explore the experiences of women who look like me, despite the popularity of intersectionality.

The results of the intersectional models were presented in Table 8.5. These models include individual characteristics, including race, gender, and the intersection between those represented by an interaction term for minority women. Interestingly, I found that these variables contributed to nearly three times more variation in 'economic security' (6.2%) than it does in other dimensions (1.9% in 'labour force participation' and 1.5% in 'marriage as a social resource'). This indicates that there are variables within the 'economic security' domain that are particularly concerned with individuals at the intersection of race and gender.

I unpicked some of the relationships between Black women and multidimensional disadvantage. When compared to White women, Black women, on average, have significantly more advantage in 'labour force participation.' Because we expect that most of an individual's income is derived from work, we would have expected to see a similar pattern between these two groups in the 'economic security' factor, particularly as income is a component of this factor. This was not the case. Black women may have more advantage in the 'labour force participation' factor, but they are less advantaged than White women in 'economic security,' on average. This highlights two issues. Firstly, the findings lend weight to some European literature that suggests that efforts to decrease social exclusion) should not rely solely on increasing the labour force participation rate (Leach *et al*., 2010). This is certainly true in this case because there are multiple forms of disadvantage uncovered in this analysis. An individual may be disadvantaged in one dimension and advantaged in other. This point is highlighted by Baca Zinn and Thornton Dill (1996) who note that some at one intersection may be on the right side of advantage in one case and on the wrong side in another. Secondly, the results highlight a need to look at the other components of economic security to uncover why Black women are still less advantaged. Previous research has suggested that Black women suffer from low rates of health insurance and low levels of income, despite high labour force participation rates (DuMonthier *et al*., 2017), both of which are components of the 'economic security' factor. Additionally, this finding reiterates a point I acknowledged earlier. It may be more important to discuss these dimensions of disadvantage separately. We have to address the various issues that comprise each dimension.

In adopting an intersectional framework to explore individual characteristics' association with disadvantage, relationships have been found that a focus on singular sociodemographic characteristics was unable to uncover. It further highlights that different forms of disadvantage are experienced differently by different groups of the population (Kabeer, 2000).

## 10.4 Was there variation in multidimensional disadvantage across the United States and the District of Columbia?

The analysis in Chapter 9 explored contextual heterogeneity in disadvantage via the use of multilevel modelling. The premise behind utilising a multilevel model is that individuals are shaped by and respond to the context in which they live (DiPrete and Forristal, 1994; Saltkjel *et al*., 2013). The results of twelve multilevel models uncovered that there is state-level variation in each factor of disadvantage. The analysis indicated that there was still variation across states whilst controlling for individual sociodemographic characteristics. In addition, I also found that the effect of gender, its relationship to each dimension, varied across states. This highlights the importance of place in understanding disadvantage.

The results in this analysis suggest that there is some variation across states for each dimension. However, the variation is quite small. A possible explanation for this is that the state is too big, as a unit, to analyse disadvantage. Within the state, there are smaller levels of government that may further influence (or be greater associated with) individual (dis)advantage in each dimension. It was beyond the scope of this thesis to include another level of analysis to the various multilevel models. Possibilities for future research here is discussed in section 10.6.

## 10.5 Policy Implications

This section focuses on the policy implications that have presented itself from the analysis in Chapter 7 and 8. I focus here because the policy recommendations offered can be addressed at the state level (the focus of Chapter 9). In Chapter 2 (section 2.5), I highlighted that most of the provision of services and policy formulation that can prevent and ease the persistence of disadvantage occur at the state level. Subsequently, the analysis suggests the following policy recommendations: 1) continue relevant income-based provisions such as the provision of food stamps that would enhance economic security, 2) continue the provision of health insurance coverage in the United States, which is currently available via the Affordable Care Act, and 3) encourage educational attainment. As disadvantage in the United States focuses on a lack of sufficient income, policy is focused on economically based interventions. However, income-based policies are not sufficient for tackling multiple forms of disadvantage uncovered in this analysis. Subsequently, I offer policy recommendations that recognise income's importance but considers the importance of other facets of disadvantage uncovered in this thesis.

I suggest that there be a continued focus on income-based policies because income poverty had a strong negative association with the 'economic security' factor. I offer two conditions for this policy recommendation. Firstly, these policies are recommended to extend aid to individuals who are 250% of the federal policy line as opposed to 100%, which is the official poverty line. Recall from Chapter 5 that I do not include the official poverty line as my indicator for income poverty status. Like others (Brady *et al*., 2013) who acknowledge the limitations of the official measure, I use a more relative income measure. I measure income poverty at 250% of the federal poverty line. It is done purposefully to acknowledge the individuals, families, and households who may be at 250% mark. They represent a key group who typically face economic insecurity and do not receive much-needed benefits. (Kearney *et al*., 2013). Secondly, the models presented throughout chapter 8 suggest that these types of provisions should be handled differently across the subgroups of the populations. For instance, in the age category, individuals aged between 18 and 24 had the lowest average score in 'economic security,' holding all the other characteristics constant. Enhancing their levels of economic security may be related to the third policy suggestion that recommends the encouragement of educational attainment. This could be handled via income supplements whilst in higher education.

The second policy recommendation is to continue the provision of health insurance coverage for all Ameircans. Health insurance coverage is important because it is directly related to economic security (Dhongde and Haveman, 2016; Hacker, 2011). With health insurance coverage, an individual becomes better able to handle the risks and shocks of ill health and death. In addition, the lack of health insurance is still recognised to be a notable contribution to Black women's enhanced risk of economic insecurity (DuMonthier *et al.*, 2017), despite the implementation of the Affordable Care Act (ACA) in 2010. Research has shown that the ACA has significantly narrowed racial disparities in health insurance coverage, care utilisation, and out of pocket expenses among breast cancer survivors (White-Means and Osmani, 2018). This suggests that the ACS can provide substantial benefits beyond the care for cancer to regular health care treatment. Efforts have to be made to dispel the legacy of inadequate health care that has exacerbated Black women's complex relationship with health care in the United States (Prather *et al.*, 2018). These efforts have to be handled at the state level, as they are typically responsible for services and amending the various provisions related to ACA. New models have to be developed to promote health that is in line with an understanding of race and gender-specific issues. Prather *et al.* (2018) suggest that

> ...programs designed to address individual-level (i.e., self-esteem, resilience), interpersonal-level (i.e., reducing stigma), community-level (i.e., reducing residential segregation), and importantly system-level factors (i.e., reducing unemployment) might facilitate long-term, sustainable improvements in health for the larger population of African American women. (pg.255)

Such efforts within each state would aid in ensuring that a population with historical and structural barriers to health care will be addressed. In so doing, the provision of health care would be greatly enhanced for everyone in the United States.

I offer the policy recommendation to encourage educational attainment because of its relationship with the 'economic security' dimension of (dis)advantage. There is universal agreement about the relevance of education in understanding disadvantage (Alkire *et al.*, 2010). The ability to participate in and pursue education affects a person's earnings (a component of the 'economic security' dimension) and has an impact on their quality of life. Educational attainment had the strongest positive relationship with 'economic security.' I suggest that there be efforts made to encourage equal access to higher education. This could be done by completely eradicating in-state tuition fees so that every individual has the ability to have the ability to increase the 'economic security' in the future.

These policy recommendations highlight the role of the state in understanding disadvantage. It also highlights their role in eradicating it. I have offered a few suggestions from this analysis and believe it would be a good first step in breaking down barriers of advantage across the United States.

## 10.6 Research Limitations

In this section, I outline some of the most notable limitations of this research. Most are the consequence of rationed resources, time or otherwise. These, then, could be addressed with further study.

There is recognition in the literature that income poverty and social exclusion vary across space and time (Bäckman and Nilsson, 2011; Bane and Ellwood, 1983; Barnes, 2005; Bossert et al., 2007; Hojman and Kast, 2009; Hunting et al., 2015; Rank, 2006). Subsequently, longitudinal data are often used to capture these dynamics (Bynner, 2000; Oroyemi *et al.*, 2010). The data source I have chosen to use in this study is cross-sectional and does not follow a group of people. Thus, temporal ambiguity is a limitation of this study, which is inherent to cross-sectional research. While a cross-sectional data set, I still believe the ACS to be the best available American data source for this analysis. It is the only data set that has relevant economic and social variables (Dhongde and Haveman, 2016) in which to measure social exclusion at the individual level. Subsequently, the limitation of using the ACS PUMS is that I am only able to provide a bounded snapshot of the individual experience of multidimensional disadvantage in the United States at one point in time.

A second limitation regards the latent construct, social exclusion. In order to operationalise this concept, it was defined and measured using the B-SEM framework. The weakness of constructing social exclusion in this study is the reliance on existing variables in the ACS PUMS that were not collected specifically to measure this concept. Consequently, for a few of the subdomains of the B-SEM, such as political participation and crime, there were no suitable indicators available. This means that I may be missing indicators that embody important components of social exclusion that would have loaded significantly in the factor analysis in Chapter 7, such as indicators relating to neighbourhood quality and satisfaction (addressed in section 5.2.6). I am confident, however, with the indicators available in the ACS for individuals aged 18 and up, the factors are consistent during the Financial Crisis (using 2008 data) and after the Financial Crisis (using the 2015 data) under analysis in this thesis. The factor loadings for 2008 data are included in Appendix B of this thesis.

An additional limitation relates to some of the variables selected from the ACS PUMS data. Some of the variables used have distinct limitations. For instance, the variable 'food stamps recipient' was used as an indicator for the economic resources subdomain of the B-SEM. I acknowledged in Chapter 3 (section 3.4.3) that a key feature of this research that distinguishes it from other poverty research is the focus on the individual as the unit of analysis, rather than the household. Food stamps are provided at the household level in the data set. This is likely because eligibility for food stamps in the United States is determined at the household level. By using this variable as a component of 'economic resources,' I could not be 100% confident that each individual within the respective housing unit had equal use of the food stamps. Additionally, 'employment status' is utilised as an indicator for the 'economic participation' subdomain of the B-SEM. The categories of this variable, 'civilian employed,' 'unemployed,' 'in the armed forces,' and 'not in the labour force,' offer a simplistic look at employment status. Consequently, I was not able to examine various forms of underemployment such as part-time work, overqualified workers with jobs mismatched to their skills (De Jong and Madamba, 2001), or the working income poor (Brady *et al.*, 2013). I was also not able to explicitly examine individuals who are undertaking unpaid work, such as an internship, or providing unpaid care. Such limitations relate to utilising 'found data' (Connelly *et al.*, 2016), in which the data set I am using was designed explicitly for this project.

Another limitation is the incomplete information available in the ACS PUMS to analyse social exclusion for two specific populations: individuals under the age of 18 and group quarters population. As acknowledged in Chapter 3, the B-SEM is acknowledged to be theoretically relevant across the life course. However, I had insufficient information to measure social exclusion for children and individuals living in group quarters. Had I had all variables for children, for instance, I may have found that there were different factors of disadvantage whilst taking this age group into consideration. Additionally, the group quarters population includes the institutionalised, a demographic who may face specific disadvantages not captured in this analysis.

In addition, I was not able to explore within-household variation in disadvantage. After deleting observations for individuals under the age of 18, I was left with one individual per household on average. This meant I was not able to include the household as a specific level of analysis in the multilevel models presented in Chapter 9. Consequently, I missed analysing within household variation in multidimensional disadvantage. This limitation means I was unable to explore within household gender differentials in disadvantage in the United States.

Finally, my analysis had to be amended due to computing limitations. As I acknowledged in Chapter 7, I had to analyse a random subsample of the data as medium power computing could not process a data set with 2 million sample members. This issue was most pronounced for factor analysis, which is quite a rigorous method of analysis. While the random subsample is still rather large, I was not able to complete an analysis on the entire data set available.

## 10.7 Recommendations for future research

There are many potential future research projects that arise from this thesis. The following section presents those ideas.

Firstly, the analysis in Chapter 7 unveiled three factors of (dis)advantage for the United States. This, however, was found using the information for sample members over the age of 18. In these specific domains individuals with various sociodemographic characteristics had varied relationships with the dimensions of (dis)advantage. For instance, I found that women, on average, were less advantaged in every factor compared to men, holding the other characteristics in the model constant (Chapter 8). Future research may seek to explore if the factors of (dis)advantage are the same for men and women. This could be done by conducting two separate factor analyses with just male sample members and the other with female sample members. Additionally, Chapter 8 also highlighted that on average, all minorities - except for Asians – in this analysis were less advantaged than Whites in every dimension of disadvantage, on average. Subsequently, this points to a need for multiracial conceptualisations of disadvantage (Sampson and Lauritsen, 1997). In this case, a factor analysis would be conducted for all sample members with a particular race of interest. It may be found that dimensions of (dis)advantage are not consistent for all groups in the American population.

Secondly, I acknowledged in Chapter 2 that the context (place) is important in understanding the complex nature of disadvantage. Indeed, it has been recognised elsewhere in the literature (Buck, 2001; Cotter, 2002). States represent an essential role in the policy arena as they are primary governors of various policies related to the prevention of disadvantage. Chapter 9 recognised that there was state-level variation in disadvantage across the United States. However, I did not include any state-level characteristics in order to aid in explaining why this variation exists. Future research could address this by examining state-level predictors. This would allow us to explore the extent to which these types of characteristics, such as a minimum wage or the prevalence of unions, relate to disadvantage.

Next, the small amount of variation found to exist between states was quite small. This partly suggests that future research could look at lower levels within the state. Policy occurs at lower levels within the state, such as the county or city that may directly relate to the individual experience of disadvantage. Subsequently, there could be an examination of within-state variation in disadvantage. Future research could start by looking at within state variation for states located in the northeast and southern regions of United States, as these regions have ranked highest and lowest in each of the dimensions of disadvantage, as acknowledged in Chapter 9. This would likely include utilising the 3 or 5 year ACS PUMS because these include geographic areas with populations less than 65,000. Using a single year ACS PUMS meant that the size of geographic units was limited to areas with large populations: states.

With the cross-sectional nature of the ACS PUMS data, it is not possible to analyse social exclusion over an individual's life. However, it does become possible to employ a repeated cross-sectional analysis in order to explore social exclusion dynamics for subgroups of the American population. Time and computing power limitations required that looking at social exclusion across time in the United States and its relationships across groups was not possible (discussed in Chapter 7). I was able to analyse 10% of the available sample population even with the limited computing capacity. If subpopulation group dynamics were analysed across several years (for instance, since 2006 when the ACS was fully implemented to the latest year of data), we would be left with what becomes an unusually large and systematic dataset. From it, we could draw relevant conclusions with reasonable confidence about how subgroups experience disadvantage in the United States over time.

Finally, in chapters 8 and 9, the different factors of (dis)advantage were examined separately. While this approach is sufficient to examine the contextual, state-level influences on a particular factor (chapter 9) and to explore the relationship between the factors and independent sociodemographic characteristics (chapter 8), there is an underlying assumption that the dimensions of (dis)advantage are not operating in tandem. By just providing the single response multilevel models, the correlation between the factors is overlooked. This correlation exists inherently (and evidence is provided in Table 7.4). Paugam (1996), using cross-sectional survey data, found that his two forms of detachment (labour force participation and social resources via the extended family) tend to go together. As I empirically found similar dimensions of exclusion, future research could model these dimensions simultaneously via the use of multivariate response multilevel models (Goldstein *et al.*, 2009).

## 10.8 Final Words

In conclusion, this data-driven research has revealed that disadvantage in the United States is indeed multidimensional. Income is but one component of disadvantage. By applying the concept of social exclusion to big American data, this analysis lends weight to the principal argument of this thesis that disadvantage is a multidimensional phenomenon with various economic and non-economic manifestations.

Utilising the concept of intersectionality in tandem with the analysis of social exclusion, this thesis has also highlighted that individuals at the intersection of characteristics (particularly Black women) experience disadvantage differently, on average, compared to individuals with those characteristics alone (for women and minorities). This has highlighted that one size fits all approaches to understanding and subsequently conquering disadvantage do not work equally across the United States. This then suggests that policy efforts should specifically target individuals and groups in the dimensions in which they are most disadvantaged.

Lastly, this analysis has highlighted that the state plays a significant role in the prevalence of disadvantage in the United States. There are differences between the states in the dimensions of disadvantage. It suggests that the states have an essential role in decreasing the risk and prevalence of disadvantage.

Taken together, this thesis has provided substantial contributions to the literature. I have looked outside of the United States to examine a concept that is incredibly under-researched in America. Not only have I applied that concept with a theoretically defined framework, the B-SEM, and derived multiple dimensions of disadvantage, I have explored that relationship with sociodemographic characteristics that confound disadvantage and assessed state-level variation. Here, I have offered an exploratory analysis of multiple forms of disadvantage. The recommendations for future research indicate that there is much more that can be learned from this thesis in order to fully understand the nature and complexity of disadvantage in the United States.

# BIBLIOGRAPHY

Aarts, E., Verhage, M., Veenvliet, J.V., Dolan, C.V., Sluis, S. van der. (2014) A solution to dependency: using multilevel analysis to accommodate nested data. *Nature Neuroscience* 17, 491–496. https://doi.org/10.1038/nn.3648

Abraham, K.G. (2015) Why labor-force participation shrank and what it means for the American economy. *The Milken Institute Review: A Journal of Economic Policy* 17, 34.

Acemoglu, D., Johnson, S., Robinson, J. (2004) Institutions as the Fundamental Cause of Long-Run Growth (Working Paper No. 10481). National Bureau of Economic Research.

Adams, R.H., Page, J. (2001) *Holding the line: Poverty reduction in the Middle East and North Africa, 1970-2000.* Poverty Reduction Group, The World Bank, Washington DC.

Adler, N.E., Newman, K. (2002). Socioeconomic Disparities in Health: Pathways and Policies. *Health Affairs* 21, 60–76. https://doi.org/10.1377/hlthaff.21.2.60

Agee, M.D., Crocker, T.D. (2013) Operationalizing the capability approach to assessing well-being. *The Journal of Socio-Economics* 46, 80–86. https://doi.org/10.1016/j.socec.2013.07.003

Ahmedani, B.K. (2011) Mental Health Stigma: Society, Individuals, and the Profession. *Journal of Social Work Values and Ethics* 8, 4-1-4–16.

Albrecht, G.L., Devlieger, P.J., (1999) The disability paradox: high quality of life against all odds. S*ocial Science & Medicine* 48, 977–988. https://doi.org/10.1016/S0277-9536(98)00411-0

Alkire, S. (2005) Why the Capability Approach? *Journal of Human Development* 6, 115–135. https://doi.org/10.1080/146498805200034275

Alkire, S., Santos, M.E. (2010) *Acute multidimensional poverty a new index for developing countries*. University of Oxford, Poverty and Human Development Initiative, Oxford, UK.

Amato, P.R., Maynard, R.A. (2007) Decreasing Nonmarital Births and Strengthening Marriage to Reduce Poverty. *The Future of Children* 17, 117–141. https://doi.org/10.1353/foc.2007.0012

Angel, J.L., Jiménez, M.A., Angel, R.J. (2007) The Economic Consequences of Widowhood for Older Minority Women. *Gerontologist* 47, 224–234. https://doi.org/10.1093/geront/47.2.224

Atkinson, A.B. (2003) Multidimensional Deprivation: Contrasting Social Welfare and Counting Approaches. *Journal of Economic Inequality* 1, 51.

Atkinson, A.B. (1998) *Poverty in Europe*. London ; Malden, Mass. : Blackwell, 1998.

Atkinson, A.B., Hills, J. (1998). Exclusion, employment and opportunity. LSE STICERD research paper no. CASE004.

Babajanian, B., Hagen-Zanker, J. (2012) *Social protection and social exclusion: an analytical framework to assess the links*. Background Note. London: ODI.

Baca Zinn, M., Thornton Dill, B. (1996) Theorizing Difference from Multiracial Feminism. *Feminist Studies* 22, 321. https://doi.org/10.2307/3178416

Bäckman, O., Nilsson, A., 2011. Pathways to Social Exclusion—A Life-Course Study. *European Sociological Review* 27, 107–123. https://doi.org/10.1093/esr/jcp064

Bailey, N., Fahmy, E., Bradshaw, J. (2018) The multidimensional analysis of social exclusion, in: Bramley, G., Bailey, N. (Eds.), *Poverty and Social Exclusion in the UK. Volume 2: The Dimensions of Disadvantage, Studies in Poverty, Inequality and Social Exclusion.* Policy Press, Bristol, UK, pp. 311–341.

Baker, R.S., 2015. The Changing Association Among Marriage, Work, and Child Poverty in the United States, 1974–2010. *Family Relations* 77, 1166–1178. https://doi.org/10.1111/jomf.12216

Bane, M.J., Ellwood, D.T. (1983) *Slipping into and out of Poverty: The Dynamics of Spells* (Working Paper No. 1199). National Bureau of Economic Research.

Barnes, C., Mercer, G. (2005) Disability, work, and welfare challenging the social exclusion of disabled people. *Work, Employment & Society* 19, 527–545.

Barnes, M. (2005) *Social Exclusion in Great Britain: An Empirical Investigation and Comparison with the EU*. Ashgate Publishing Limited, England.

Bauer, G.R., Scheim, A.I. (2019) Methods for analytic intercategorical intersectionality in quantitative research: Discrimination as a mediator of health inequalities. *Social Science & Medicine* 226, 236–245. https://doi.org/10.1016/j.socscimed.2018.12.015

Baulch, B. (1996) Editorial: The New Poverty Agenda: A Disputed Consensus. *IDS Bulletin* 27, 1–10. https://doi.org/10.1111/j.1759-5436.1996.mp27001001.x

Beals, K.P., Peplau, L.A. (2005) Identity Support, Identity Devaluation, and Well-being among Lesbians. *Psychology of Women Quarterly* 29, 140–148.

Beavers, A.S., Lounsbury, J.W., Richards, J.K., Huck, S.W., Skolits, G.J., Esquivel, S.L. (2013) Practical considerations for using exploratory factor analysis in educational research. *Practical assessment, research & evaluation* 18.

Bebbington, A.J., Dani, A.A., Haan, A. de, Walton, M. (2008) *Inequalities and Development: Dysfunctions, Traps and Transitions, in: Institutional Pathways to Equity: Addressing Inequality Traps*. World Bank Publications, Washington D.C.

Bech, M., Gyrd-Hansen, D. (2005) Effects coding in discrete choice experiments. *Health Economics* 14, 1079–1083. https://doi.org/10.1002/hec.984

Beeghley, L. (1988) Individual and Structural Explanations of Poverty. *Population Research and Policy Review* 7, 201–222.

Berube, A., others (2012) *The continuing evolution of American poverty and its implications for community development*. Investing in What Works for America's Communities 55–71.

Besharov, D.J., Couch, K. (2009) European measures of income, poverty, and social exclusion: Recent developments and lessons for U.S. poverty measurement. *Journal of Policy Analysis and Management* 28, 713–715. https://doi.org/10.1002/pam.20467

Besley, T., Coate, S. (1992) Workfare versus Welfare: Incentive Arguments for Work Requirements in Poverty-Alleviation Programs. *The American Economic Review* 82, 249–261.

Betson, D.M., Warlick, J.L. (1998) Alternative Historical Trends in Poverty. T*he American Economic Review* 88, 348–351.

Betti, G., Verma, V. (1999) Measuring the degree of poverty in a dynamic and comparative context: a multi-dimensional approach using fuzzy set theory, in: *Proceedings*, ICCS-VI. p. 289.

Bhopal, K. (2018). *White Privilege: The Myth of a Post-Racial Society*. Policy Press, Bristol, UK.

Bland, J.M., Altman, D.G. (1997) Cronbach's alpha. *British Medical Journal* 314, 572.

Blank, R.M. (2008) Presidential address: How to improve poverty measurement in the United States. *Journal of Policy Analysis and Management* 27, 233–254. https://doi.org/10.1002/pam.20323

Blank, R.M. (2003) Selecting Among Anti-Poverty Policies: Can an Economist Be Both Critical and Caring? *Review of Social Economy* 61, 447–469. https://doi.org/10.1080/0034676032000160949

Blau, J.R., Blau, P.M. (1982) The Cost of Inequality: Metropolitan Structure and Violent Crime. *American Sociological Review* 47, 114–129. https://doi.org/10.2307/2095046

Bloom, D.E., Grenier, G. (2009) Language, employment, and earnings in fhe United States: Spanish-English differentials from 1970 to 1990. I*nternational Journal of the Sociology of Language* 121, 45–68. https://doi.org/10.1515/ijsl.1996.121.45

Bolker, B.M., Brooks, M.E., Clark, C.J., Geange, S.W., Poulsen, J.R., Stevens, M.H.H., White, J.-S.S. (2009) Generalized linear mixed models: a practical guide for ecology and evolution. T*rends in Ecology & Evolution* 24, 127–135. https://doi.org/10.1016/j.tree.2008.10.008

Bonilla-Silva, E. (1997) Rethinking Racism: Toward a Structural Interpretation. *American Sociological Review* 62, 465. https://doi.org/10.2307/2657316

Boniwell, I., Osin, E.N., Renton, A. (2015) Internet Access at Home and its Relationship to Well-being in Deprived Areas of London. *The Open Psychology Journal* 8, 44–53. https://doi.org/10.2174/1874350101508010044

Bossert, W., Chakravarty, S.R., D'Ambrosio, C. (2012) Multidimensional Poverty and Material Deprivation with Discrete Data. *Review of Income and Wealth* 59, 29–43. https://doi.org/10.1111/j.1475-4991.2012.00519.x

Bossert, W., D'ambrosio, C. (2013) Measuring Economic Insecurity. *International Economic Review* 54, 1017–1030. https://doi.org/10.1111/iere.12026

Bossert, W., D'ambrosio, C., Peragine, V. (2007) Deprivation and Social Exclusion. *Economica* 74, 777–803. https://doi.org/10.1111/j.1468-0335.2006.00572.x

Boudreaux, M., Ziegenfuss, J.Y., Graven, P., Davern, M., Blewett, L.A. (2011) Counting Uninsurance and Means-Tested Coverage in the American Community Survey: A Comparison to the Current Population Survey: Counting Uninsurance and Means-Tested Coverage. *Health Services Research* 46, 210–231. https://doi.org/10.1111/j.1475-6773.2010.01193.x

Bowleg, L. (2012) The Problem with the Phrase Women and Minorities: Intersectionality—an Important Theoretical Framework for Public Health. *American Journal of Public Health* 102, 1267–1273. https://doi.org/10.2105/AJPH.2012.300750

Bowleg, L. (2008) When Black + Lesbian + Woman ≠ Black Lesbian Woman: The Methodological Challenges of Qualitative and Quantitative Intersectionality Research. *Sex Roles* 59, 312–325. https://doi.org/10.1007/s11199-008-9400-z

Brady, D. (2009) *Rich democracies, poor people: how politics explain poverty*. New York ; Oxford : Oxford University Press, c2009.

Brady, D. (2004) Reconsidering the Divergence between Elderly, Child, and Overall Poverty. *Research on Aging* 26, 487–510. https://doi.org/10.1177/0164027504266587

Brady, D. (2003). Rethinking the Sociological Measurement of Poverty. *Social Forces* 81, 715–751. https://doi.org/10.1353/sof.2003.0025

Brady, D., Baker, R.S., Finnigan, R. (2013) When Unionization Disappears State-Level Unionization and Working Poverty in the United States. *American Sociological Review* 78, 872–896. https://doi.org/10.1177/0003122413501859

Brady, D., Burroway, R. (2012) Targeting, Universalism, and Single-Mother Poverty: A Multilevel Analysis Across 18 Affluent Democracies. *Demography* 49, 719–746. https://doi.org/10.1007/s13524-012-0094-z

Brady, D., Fullerton, A.S., Cross, J.M. (2009) Putting poverty in political context: a multi-level analysis of adult poverty across 18 affluent democracies. *Social Forces* 88, 271–299.

Bramley, G., Bailey, N. (Eds.) (2018) *Poverty and Social Exclusion in the UK: Vol. 2: The Dimensions of Disadvantage*. Policy Press.

Brooks-Gunn, J., Duncan, G.J. (1997) The Effects of Poverty on Children. *The Future of Children* 7, 55–71. https://doi.org/10.2307/1602387

Brown, D.L., Hirschl, T.A. (1995) Household Poverty in Rural and Metropolitan-Core Areas of the United States. *Rural Sociology* 60, 44–66.

Brown, P. (1995) Cultural Capital and Social Exclusion: Some Observations on Recent Trends in Education, Employment and the Labour Market. Work, *Employment & Society* 9, 29–51. https://doi.org/10.1177/095001709500900103

Buck, N. (2001) Identifying Neighbourhood Effects on Social Exclusion. *Urban Studies* 38, 2251–2275. https://doi.org/10.1080/00420980120087153

Bumpass, L.L., Sweet, J.A., Cherlin, A. (1991) The Role of Cohabitation in Declining Rates of Marriage. *Journal of Marriage and Family* 53, 913–927. https://doi.org/10.2307/352997

Burchardt, T., Hick, R. (2018) Inequality, Advantage and the Capability Approach. *Journal of Human Development and Capabilities* 19, 38–52. https://doi.org/10.1080/19452829.2017.1395396

Burchardt, T., Le Grand, J., Piachaud, D. (2002) Introduction, in: Hills, J., Le Grand, J., Piachaud, D. (Eds.), *Understanding Social Exclusion*. Oxford University Press, Oxford, UK, pp. 1–12.

Burchardt, T., Le Grand, J., Piachaud, D. (1999) Social exclusion in Britain 1991—1995. S*ocial Policy & Administration* 33, 227–244.

Bureau of Labor Statistics (2018) *Consumer Price Index - March 2018*. United States Department of Labor, Bureau of Labor Statistics, Washington, United States, Washington.

Bureau of Labor Statistics (2016) C*onsumer Expenditures in 2015* (No. USDL-16-1768). United States Department of Labor, Bureau of Labor Statistics, Washington, D.C.

Burkhauser, R.V., Daly, M.C., Ziebarth, N.R. (2016) Protecting working-age people with disabilities: experiences of four industrialized nations. *Journal for Labour Market Research* 49, 367–386. https://doi.org/10.1007/s12651-016-0215-z

Burns, P., Gimpel, J.G. (2000) Economic Insecurity, Prejudicial Stereotypes, and Public Opinion on Immigration *Policy. Political Science Quarterly* 115, 201–225. https://doi.org/10.2307/2657900

Bwire, A., Rodríguez-Pose, A. (2003) T*he economic (in)efficiency of devolution*. Department of Geography and Environment, LSE, London.

Bynner, J. (2000) *Risks and Outcomes of Social exclusion: insights from longitudinal data*. Institute of Education University of London, London.

Caiola, C., Docherty, S., Relf, M., Barroso, J. (2014) Using an Intersectional Approach to Study the Impact of Social Determinants of Health for African-American Mothers Living with HIV. *Advancing Nursing Science* 37, 287–298. https://doi.org/10.1097/ANS.0000000000000046

Cancian, M., Danziger, S. (2009) *Changing Poverty, Changing Policie*s. Russell Sage Foundation.

Carle, A.C. (2009) Fitting multilevel models in complex survey data with design weights: Recommendations. *BMC Medical Research Methodology* 9, 49. https://doi.org/10.1186/1471-2288-9-49

Catalano, R. (1991) The health effects of economic insecurity. *American Journal of Public Health* 81, 1148–1152. https://doi.org/10.2105/AJPH.81.9.1148

Cellini, S.R., McKernan, S.-M., Ratcliffe, C. (2008) The dynamics of poverty in the United States: A review of data, methods, and findings. *Journal of Policy Analysis and Management* 27, 577–605. https://doi.org/10.1002/pam.20337

Chakravarty, S.R. (2006) An axiomatic approach to multidimensional poverty measurement via fuzzy sets, in: F*uzzy Set Approach to Multidimensional Poverty Measurement*. Springer, pp. 49–72.

Chambers, R. (1995) Poverty and livelihoods: whose reality counts? *Environment and Urbanization* 7, 32.

Chant, S. (2003) *Female Household Headship and the Feminisation of Poverty: Facts, Fictions and Forward Strategies*. LSE Research Online, London.

Chen, H., Chiang, R.H.L., Storey, V.C., 2012. Business Intelligence and Analytics: From Big Data to Big Impact. *MIS Quarterly* 36, 1165. https://doi.org/10.2307/41703503

Chen, K.-M., Wang, T.-M. (2015) Determinants of Poverty Status in Taiwan: A Multilevel Approach. *Social Indicators Research* 123, 371–389. https://doi.org/10.1007/s11205-014-0741-4

Cheng, H.G., Phillips, M.R. (2014) Secondary analysis of existing data: opportunities and implementation. *Shanghai Archives of Psychiatry* 26, 371–375. https://doi.org/10.11919/j.issn.1002-0829.214171

Cherlin, A.J. (2004) The deinstitutionalization of American marriage. *Journal of Marriage and Family* 66, 848–861. https://doi.org/10.1111/j.0022-2445.2004.00058.x

Chiswick, B.R., Miller, P.W. (2002) Immigrant earnings: Language skills, linguistic concentrations and the business cycle. *Journal of Population Economics* 15, 31–57. https://doi.org/10.1007/PL00003838

Chiswick, B.R., Miller, P.W. (1995) The Endogeneity between Language and Earnings: International Analyses. *Journal of Labour Economics* 13, 246–288. https://doi.org/10.1086/298374

Cho, S., Crenshaw, K.W., McCall, L. (2013) Toward a Field of Intersectionality Studies: Theory, Applications, and Praxis. *Signs: Journal of Women in Culture and Society* 38, 785–810. https://doi.org/10.1086/669608

Citro, C.F., Michael, R.T. (1995) *Measuring Poverty: A New Approach*. National Academy Press, Washington D.C.

Cleophas, T.J., Zwinderman, A.H. (2018) Continuous Outcome Regressions, in: *Regression Analysis in Medical Research*. Springer, Cham, pp. 1–39. https://doi.org/10.1007/978-3-319-71937-5_1

Coleman, J.S. (1988) Social capital in the creation of human capital. *American Journal of Sociology* 94, S95–S120.

Conley, D. (2005) Poverty and life chances: The conceptualization and study of the poor. *The SAGE Handbook of Sociology* 327.

Connelly, R., Playford, C.J., Gayle, V., Dibben, C. (2016) The role of administrative data in the big data revolution in social science research. *Social Science Research*, Special issue on Big Data in the Social Sciences 59, 1–12. https://doi.org/10.1016/j.ssresearch.2016.04.015

Cook, J.T., Black, M., Chilton, M., Cutts, D., Ettinger de Cuba, S., Heeren, T.C., Rose-Jacobs, R., Sandel, M., Casey, P.H., Coleman, S., Weiss, I., Frank, D.A. (2013) Are Food Insecurity's Health Impacts Underestimated in the U.S. Population? Marginal Food Security Also Predicts Adverse Health Outcomes in Young U.S. Children and Mothers123. *Advances in Nutrition* 4, 51–61. https://doi.org/10.3945/an.112.003228

Cook, J.T., Frank, D.A. (2008) Food Security, Poverty, and Human Development in the United States. *Annals of the New York Academy of Sciences* 1136, 193–209. https://doi.org/10.1196/annals.1425.001

Cornwell, E.Y., Waite, L.J. (2009) Social Disconnectedness, Perceived Isolation, and Health among Older Adults. *Journal of Health and Social Behaviour* 50, 31–48.

Costello, A.B., Osborne, J.W. (2005) Best Practices in Exploratory Factor Analysis: Four Recommendations for Getting the Most from Your Analysis. Practical Assessment, *Research & Evaluation* 10, 1–9.

Cotter, D.A. (2002) Poor people in poor places: local opportunity structures and household poverty. *Rural Sociology* 534.

Craney, T.A., Surles, J.G. (2002) Model-Dependent Variance Inflation Factor Cutoff Values. *Quality Engineering* 14, 391–403. https://doi.org/10.1081/QEN-120001878

Craven, B.D., Islam, S.M.N. (2011) Ordinaray Least-Squares Regression, in: Moutinho, L., Hutcheson, G.D. (Eds.), *The SAGE Dictionary of Quantitative Management Research*. SAGE, pp. 224–227.

Crenshaw, K.W. (2014) Beyond Racism and Misogyny: Black feminism and 2 Live Crew, in: Meyers, D.T. (Ed.), *Feminist Social Thought: A Reader*. Routledge, New York, pp. 245–263.

Crenshaw, K.W. (1994) Mapping the Margins: Intersectionality, Identity Politics, and Violence Against Women of Color, in: Fineman, M.A., Mykitiuk, R. (Eds.), *The Public Nature of Private Violence*. Routledge, New York, pp. 93–118.

Crenshaw, K., (1991). Mapping the Margins: Intersectionality, Identity Politics, and Violence against Women of Colour. *Stanford Law Review* 43, 1241–1299. https://doi.org/10.2307/1229039

Crenshaw, K.W. (1989) Demarginalizing the Intersection of Race and Sex: A Black Feminist Critique of Antidiscrimination Doctrine, Feminist Theory, and Antiracist Politics. *University of Chicago Legal Forum* 1989, 57–80. https://doi.org/10.4324/9780429500480-5

Crettaz, E., Bonoli, G. (2010) Why Are Some Workers Poor? The Mechanisms that Produce Working Poverty in a Comparative Perspective (Working Paper). The University of Edinburgh.

Cronbach, L.J. (1951) Coefficient alpha and the internal structure of tests. *Psychometrika* 16, 297–334.

Crous, G., Bradshaw, J. (2017) Child social exclusion. *Children and Youth Services Review* 80, 129–139. https://doi.org/10.1016/j.childyouth.2017.06.062

Cullen, R. (2001) Addressing the digital divide. *Online Information Review* 25, 311–320. https://doi.org/10.1108/14684520110410517

Cunradi, C.B., Caetano, R., Clark, C., Schafer, J. (2000) Neighbourhood poverty as a predictor of intimate partner violence among White, Black, and Hispanic couples in the United States: a multilevel analysis. *Annals of Epidemiology* 10, 297–308.

Cusworth, L., Bradshaw, J., Coles, B., Chzhen, Y. (2009) *Understanding the risks of social exclusion across the life course: Youth and Young Adulthood.* UK Cabinet Office, Social Exclusion Task Force.

Cutrona, C. (1996) What is Social Support and What Makes You Think You Have it?, in: *Social Support in Couples: Marriage as a Resource in Times of Stress*. SAGE Publications, Inc., Thousand Oaks, pp. 1–17. https://doi.org/10.4135/9781483327563

Dalaker, J. (2017) *The Supplemental Poverty Measure: Its Core Concepts, Development, and Use* (Congressional Reserch Service Report No. 7–5700). Congressional Research Service, Washington D.C.

Dannefer, D. (2003) Cumulative Advantage/Disadvantage and the Life Course: Cross-Fertilizing Age and Social Science Theory. J*ournal of Gerontology Series B Psychological Sciences and Social Sciences* 58, S327–S337. https://doi.org/10.1093/geronb/58.6.S327

Davern, M., Quinn, B.C., Kenney, G.M., Blewett, L.A. (2009) The American Community Survey and Health Insurance Coverage Estimates: Possibilities and Challenges for Health Policy Researchers. *Health Services Research* 44, 593–605. https://doi.org/10.1111/j.1475-6773.2008.00921.x

Davies, J.S. (2005) The Social Exclusion Debate: Strategies, controversies and dilemmas. *Policy Studies* 26, 3–27. https://doi.org/10.1080/01442870500041561

Davis, K. (2008) Intersectionality as buzzword: A sociology of science perspective on what makes a feminist theory successful. *Feminist Theory* 9, 67–85. https://doi.org/10.1177/1464700108086364

de Haan, A. (2000) Social Exclusion: Enriching the understanding of deprivation. Studies in Social and *Political Thought* 2, 22–40.

de Haan, A. (1999) *Social Exclusion*. Department for International Development, London.

De Jong, G.F., Madamba, A.B. (2001) A Double Disadvantage? Minority Group, Immigrant Status, and Underemployment in the United States. *Social Science Quarterly* 82, 117–130. https://doi.org/10.1111/0038-4941.00011

de Winter, J.C.F., Dodou, D. (2012) Factor recovery by principal axis factoring and maximum likelihood factor analysis as a function of factor pattern and sample size. J*ournal of Applied Statistics* 39, 695–710. https://doi.org/10.1080/02664763.2011.610445

Dean, H. (2016) Poverty and Social Exclusion, in: Platt, L., Dean, H. (Eds.), *Social Advantage and Disadvantage*. Oxford University Press, Oxford, UK, pp. 3–24.

Dermott, E., Main, G. (Eds.) (2017) *Poverty and Social Exclusion in the UK: Vol 1: The Nature and Extent of the Problem.* Policy Press.

Devlieger, I., Mayer, A., Rosseel, Y. (2016) Hypothesis Testing Using Factor Score Regression: A Comparison of Four Methods. *Educational and Psychological Measurement* 76, 741–770. https://doi.org/10.1177/0013164415607618

Dewilde, C. (2003) A life-course perspective on social exclusion and poverty. *British Journal of Sociology* 54, 109–128. https://doi.org/10.1080/0007131032000045923

Dhongde, S., Haveman, R. (2016) Multi-Dimensional Deprivation in the U.S. *Social Indicators Research*. https://doi.org/10.1007/s11205-016-1379-1

Dieleman, J.L., Templin, T. (2014) Random-Effects, Fixed-Effects and the within-between Specification for Clustered Data in Observational Health Studies: A Simulation Study. P*LoS One* 9. https://doi.org/10.1371/journal.pone.0110257

Diener, E., Suh, E. (1997) Measuring quality of life: Economic, social, and subjective indicators. *Social Indicators Research* 40, 189–216.

Dill, B.T. (1983) Race, Class, and Gender: Prospects for an All-Inclusive Sisterhood. *Feminist Studies* 9, 131–150. https://doi.org/10.2307/3177687

DiMaggio, P., Hargittai, E. (2001) *From the 'digital divide 'to 'digital inequality': Studying Internet use as penetration increases.* Princeton: Centre for Arts and Cultural Policy Studies, Woodrow Wilson School, Princeton University 4, 4–2.

DiMaggio, P., Mohr, J. (1985) Cultural Capital, Educational Attainment, and Marital Selection. *American Journal of Sociology* 90, 1231–1261.

DiPrete, T.A., Forristal, J.D. (1994) Multilevel Models: Methods and Substance. *Annual Review of Sociology* 20, 331–357.

Disatnik, D., Sivan, L. (2016) The multicollinearity illusion in moderated regression analysis. *Marketing Letters* 27, 403–408. https://doi.org/10.1007/s11002-014-9339-5

DiStefano, C., Zhu, M., Mindrila, D. (2009) Understanding and using factor scores: Considerations for the applied researcher. *Practical Assessment, Research & Evaluation* 14, 1–11.

Dominitz, J., Manski, C.F. (1996) *Perceptions of Economic Insecurity: Evidence from the Survey of Economic Expectations* (Working Paper No. 5690). National Bureau of Economic Research. https://doi.org/10.3386/w5690

Donenfeld, M.J. (1940) The Meaning of Social Resources. *Journal of Educational Sociology* 13, 560. https://doi.org/10.2307/2261721

Doornik, J.A., Hansen, H. (2008) An Omnibus Test for Univariate and Multivariate Normality. *Oxford Bulletin of Economics and Statistics* 70, 927–939. https://doi.org/10.1111/j.1468-0084.2008.00537.x

Doyal, L., Gough, I. (1991) The Grammar of 'Need,' in: Doyal, L., Gough, I. (Eds.), *A Theory of Human Need*. Macmillan Education UK, London, pp. 35–45. https://doi.org/10.1007/978-1-349-21500-3_4

Doyle, O., Pecukonis, E., Harrington, D. (2011) The Nurturant Fathering Scale: A Confirmatory Factor Analysis with an African American Sample of College Students. R*esearch on Social Work Practice* 21, 319–327. https://doi.org/10.1177/1049731510377635

Drake, R.E., Skinner, J.S., Bond, G.R., Goldman, H.H. (2009) Social Security and Mental Illness: Reducing Disability with Supported Employment. *Health Affairs* (Millwood) 28, 761–770. https://doi.org/10.1377/hlthaff.28.3.761

Dubrow, J.K. (2008) How Can We Account for Intersectionality in Quantitative Analysis of Survey Data? Empirical Illustration for Central and Eastern Europe. ASK. *Research & Methods* 85–100.

DuMonthier, A., Childers, C., Milli, J. (2017) The Status of Black Women in the United States, The Status of Women in the States. The Institute for Women's Policy Research, Washington DC.

Duncan, C.M., Tickamyer, A.R., 1988. Poverty research and policy for rural America. T*he American Sociologist* 19, 243–259. https://doi.org/10.1007/BF02691983

Eager, C., Roy, J. (2017) Mixed Effects Models are Sometimes Terrible. *arXiv e-prints*.

Easley, J. (2018) Spatial mismatch beyond black and white: Levels and determinants of job access among Asian and Hispanic subpopulations. *Urban Studies* 55, 1800–1820. https://doi.org/10.1177/0042098017696254

Egumenovska, K. (2012) Black-boxing political intersectionality: on 'othering' that deprives individuals of their citizenship. *Social Identities* 18, 679–694. https://doi.org/10.1080/13504630.2012.708998

Elliott, J.R., Joyce, M.S. (2004) The Effects of Race and Family Structure on Women's Spatial Relationship to the Labour Market. *Sociological Inquiry* 74, 411–435. https://doi.org/10.1111/j.1475-682X.2004.00098.x

Ellwood, D.T. (2000) Anti-poverty policy for families in the next century: From welfare to work–and worries. *The Journal of Economic Perspectives* 187–198.

Engberg-Pedersen, L., Ravnborg, H.M. (2010). *Conceptualisations of poverty*. Danish Institute for International Studies, Copenhagen.

Engle, J.C. (2013) Promoting the general welfare: legal reform to lift women and children in the United States out of poverty. *Journal of Gender, Race and Justice* 16.

Fabrigar, L.R., Wegener, D.T. (2011) *Exploratory Factor Analysis*. Oxford University Press. https://doi.org/10.1093/acprof:osobl/9780199734177.001.0001

Ferguson, E., Cox, T. (1993) Exploratory Factor Analysis: A Users' Guide. I*nternational Journal of Selection and Assessment* 1, 84–94. https://doi.org/10.1111/j.1468-2389.1993.tb00092.x

Fernandez, R.M., Su, C. (2004) Space in the study of labour markets. *Annual Review of Sociology* 30, 545.

Ferraro, K.F., Shippee, T.P. (2009) Aging and Cumulative Inequality: How Does Inequality Get Under the Skin? *Gerontologist* 49, 333–343. https://doi.org/10.1093/geront/gnp034

Field, A.P. (2009) *Discovering statistics using SPSS: (and sex, drugs and rock "n" roll)*, 3rd ed. SAGE Publications, Los Angeles.

Field, A.P., Wright, D.B. (2011) A Primer on Using Multilevel Models in Clinical and Experimental Psychopathology Research. *Journal of Experimental Psychopathology* 2, jep.013711. https://doi.org/10.5127/jep.013711

Fisher, G.M. (1997) T*he Development of the Orshansky Poverty Thresholds and Their Subsequent History as the Official U.S. Poverty Measure*. US Census Bureau, Washington D.C.

Fisher, G.M. (1993) *From Hunter to Orshansky: An Overview of (Unofficial) Poverty Lines in the United States from 1904 to 1965*. Presented at the Fifteenth Annual Research Conference of the Association for Public Policy Analysis and Management, Washington, D.C.

Fisher, G.M. (1992) Poverty Guidelines for 1992. *Social Security Bulletin* 55, 43–46.

Floridi, L. (2012) Big Data and Their Epistemological Challenge. *Philosophy & Technology* 25, 435–437.

Floyd, F.J., Widaman, K.F. (1995) Factor analysis in the development and refinement of clinical assessment instruments. *Psychological Assessment* 7, 286–299. https://doi.org/10.1037/1040-3590.7.3.286

Foster, J. (2000) Social Exclusion, Crime and Drugs. *Drugs: Education, Prevention and Policy* 7, 317–330. https://doi.org/10.1080/dep.7.4.317.330

Fox, L., Wimer, C., Garfinkel, I., Kaushal, N., Waldfogel, J. (2015) Waging War on Poverty: Poverty Trends Using a Historical Supplemental Poverty Measure. J*ournal of Policy Analysis and Management* 34, 567–592. https://doi.org/10.1002/pam.21833

Franklin, S.B., Gibson, D.J., Robertson, P.A., Pohlmann, J.T., Fralish, J.S. (1995) Parallel analysis: a method for determining significant principal components. *Journal of Vegetation Science* 6, 99–106.

Fremstad, S. (2008) *Measuring Poverty and Economic Inclusion: The Current Poverty Measure, the NAS Alternative, and the Case for a Truly New Approach*. December. Washington, DC: Center for Economic and Policy Research.

Furstenberg, F.F., Kaplan, S.B. (2004) Social Capital and the Family, in: Scott, J., Treas, J., Richards, M. (Eds.), *The Blackwell Companion to the Sociology of Families*. Blackwell Publishing Ltd, pp. 218–232. https://doi.org/10.1002/9780470999004.ch13

Galabuzi, G.-E. (2016) Social Exclusion, in: Raphael, D. (Ed.), S*ocial Determinants of Health: Canadian Perspectives*. Canadian Scholars' Press.

Galbraith, J.K. (2000) *Created Unequal: The Crisis in American Pay*, 1st edition. ed. University Of Chicago Press, Chicago.

Gallagher, M.W., Brown, T.A. (2013) Introduction to Confirmatory Factor Analysis and Structural Equation Modelling, in: Teo, T. (Ed.), *Handbook of Quantitative Methods for Educational Research*. Sense Publishers, Rotterdam. https://doi.org/10.1007/978-94-6209-404-8

Gandomi, A., Haider, M. (2015) Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management* 35, 137–144. https://doi.org/10.1016/j.ijinfomgt.2014.10.007

Gangl, M., Ziefle, A. (2009) Motherhood, Labor Force Behavior, and Women's Careers: An Empirical Assessment of the Wage Penalty for Motherhood in Britain, Germany, and the United States. *Demography* 46, 341–369.

Garcia, J.A. (1981) Political Integration of Mexican Immigrants: Explorations into the Naturalization Process. *International Migration Review* 15, 608–625. https://doi.org/10.1177/019791838101500401

Garmezy, N. (1993) Children in Poverty: Resilience Despite Risk. *Psychiatry* 56, 127–136. https://doi.org/10.1080/00332747.1993.11024627

Garrison, M. (2005) *Marriage: Essential or Expendable?* (SSRN Scholarly Paper No. ID 829825). Social Science Research Network, Rochester, NY.

Gee, G.C., Morey, B.N., Walsemann, K.M., Ro, A., Takeuchi, D.T. (2016) Citizenship as Privilege and Social Identity: Implications for Psychological Distress. *American Behavioural Scientist* 60, 680–704. https://doi.org/10.1177/0002764216632834

George, D., Mallery, P. (2003) *SPSS for Windows Step by Step: A Simple Guide and Reference*, 4th ed. Allyn & Bacon, Boston.

Getz, V.K. (1985) In-Kind Benefits Help Reduce Poverty. *Rural Development Perspectives* 38–40.

Geva, D. (2011) Not Just Maternalism: Marriage and Fatherhood in American Welfare Policy. *Social Politics: International Studies in Gender, State & Society* 18, 24–51. https://doi.org/10.1093/sp/jxr003

Gilens, M. (1999) *Why Americans hate welfare: race, media, and the politics of antipoverty policy, Studies in communication, media, and public opinion*. Chicago; London: University of Chicago Press, 1999.

Ginther, D.K., Zavodny, M. (2001) Is the male marriage premium due to selection? The effect of shotgun weddings on the return to marriage. *Journal of Population Economics* 14, 313–328. https://doi.org/10.1007/s001480000058

Giraud, G., Renouard, C., L'Huillier, H., De La Martinière, R., Sutter, C. (2013) *Relational capability: A multidimensional approach.*

Glaser, D., Hastings, R.H. (2011) An Introduction to Multilevel Modelling for Anaesthesiologists *Anaesthesia & Analgesia* 113, 877–887. https://doi.org/10.1213/ANE.0b013e3182198a01

Glennerster, H. (2002) United States Poverty Studies and Poverty Measurement: The Past Twenty-Five Years. *Social Service Review* 76, 83–107. https://doi.org/10.1086/324609

Goldstein, H., Carpenter, J., Kenward, M.G., Levin, K.A. (2009) Multilevel models with multivariate mixed response types. S*tatistical Modelling: An International Journal* 9, 173–197. https://doi.org/10.1177/1471082X0800900301

Gordon, David. (2000) *Poverty and social exclusion in Britain*. Jospeh Rowntree Foundation, York, UK.

Gornick, J.C., Jäntti, M. (2010) *Women, poverty, and social policy regimes: A cross-national analysis* (Working Paper No. 534). LIS Working Paper Series.

Gould, E.D., Paserman, M.D. (2003) Waiting for Mr. Right: rising inequality and declining marriage rates. *Journal of Urban Economics* 53, 257–281. https://doi.org/10.1016/S0094-1190(02)00518-1

Gould, E.D., Weinberg, B.A., Mustard, D.B. (2002) Crime Rates and Local Labor Market Opportunities in the United States: 1979 - 1997. T*he Review of Economics and Statistics* 84, 45–61.

Gove, W.R., Hughes, M., Style, C.B. (1983) Does Marriage Have Positive Effects on the Psychological Well-Being of the Individual? *Journal of Health and Social Behavior* 24, 122–131. https://doi.org/10.2307/2136639

Groves, R.M. (2012) *The Pros and Cons of Making the Census Bureau's American Community Survey Voluntary*. US Census Bureau, Washington D.C.

Grzywacz, J.G., Fuqua, J. (2000) The Social Ecology of Health: Leverage Points and Linkages. *Behavioral Medicine* 26, 101–115. https://doi.org/10.1080/08964280009595758

Gubernskaya, Z., Bean, F.D., Van Hook, J. (2013) Why Does Naturalization Matter for the Health of Older Immigrants in the United States? *Journal of Health and Social Behavior*; Washington 54, 516.

Gurka, M.J., Edwards, L.J., Muller, K.E. (2011) Avoiding bias in mixed model inference for fixed effects. *Statistics in Medicine* 30, 2696–2707. https://doi.org/10.1002/sim.4293

Haber, A., 1966. Poverty Budgets: How Much Is Enough? *Asia Pacific Journal of Human Resources* 1, 5–22. https://doi.org/10.1177/103841116600100302

Hacker, J.S. (2011) *Understanding Economic Insecurity: The Downward Spiral of the Middle Class*. Boston Federal Reserve.

Hacker, J.S. (2007) The New Economic Insecurity - And What Can Be Done about It. *Harvard Law & Policy Review* 1, 111–126.

Hacker, J.S. (2006) *The Great Risk Shift: The New Economic Insecurity and the Decline of the American Dream*. Oxford University Press.

Hacker, J.S., Huber, G.A., Nichols, A., Rehm, P., Schlesinger, M., Valletta, R., Craig, S. (2014) The Economic Security Index: A New Measure for Research and Policy Analysis. *Review of Income and Wealth 60*, S5–S32. https://doi.org/10.1111/roiw.12053

Hair Jr., J.F., Black, W.C., Babin, B.J., Anderson, R.E. (2010) *Multivariate Data Analysis*, 7th ed. Pearson, Harlow, Essex.

Haley, J.D., Rejda, G.E. (2001) Measuring the causes of economic insecurity: Construction of a composite index. *Journal of Financial Service Professionals;* Bryn Mawr 55, 103–111.

Hall, R.J. (2015) Feminist Strategies to End Violence Against Women. *The Oxford Handbook of Transnational Feminist Movements.* ttps://doi.org/10.1093/oxfordhb/9780199943494.013.005

Hampton, K.N. (2010) Internet Use and the Concentration of Disadvantage: Glocalization and the Urban Underclass. *American Behavioral Scientist* 53, 1111–1132. https://doi.org/10.1177/0002764209356244

Hancock, A.-M. (2007) When Multiplication Doesn't Equal Quick Addition: Examining Intersectionality as a Research Paradigm. *Perspectives on Politics* 5. https://doi.org/10.1017/S1537592707070065

Hand, D.J. (2018) Statistical challenges of administrative and transaction data. *Journal of Royal Statistical Society* 181, 1–24.

Hankivsky, O., Christoffersen, A. (2008) I*ntersectionality and the Determinants of Health: a Canadian Perspective*. https://doi.org/10.1080/09581590802294296

Hankivsky, O., Grace, D., Hunting, G., Ferlatte, O., Clark, N., Fridkin, A., Giesbrecht, M., Rudrum, S., Laviolette, T. (2012) Intersectionality-Based Policy Analysis, in: Hankivsky, O. (Ed.), A*n Intersectionality-Based Policy Analysis Framework*. Institute for Intersectionality Research & Policy, Simon Fraser University, Vancouver, BC, pp. 33–45.

Hanna, D., White, R., Lyons, K., McParland, M.J., Shannon, C., Mulholland, C. (2011) The structure of the Beck Hopelessness Scale: A confirmatory factor analysis in UK students. *Personality and Individual Differences* 51, 17–22. https://doi.org/10.1016/j.paid.2011.03.001

Hargittai, E. (2003) The Digital Divide and What to do About It, in: Jones, D.C. (Ed.), *New Economy Handbook*. pp. 821–839.

Harrison, X.A., Donaldson, L., Correa-Cano, M.E., Evans, J., Fisher, D.N., Goodwin, C.E.D., Robinson, B.S., Hodgson, D.J., Inger, R. (2018). A brief introduction to mixed effects modelling and multi-model inference in ecology. *PeerJ* 6. https://doi.org/10.7717/peerj.4794

Hatch, A. (2017) Saying "I Don't" to Matrimony: An Investigation of Why Long-Term Heterosexual Cohabitors Choose Not to Marry. *Journal of Family Issues* 38, 1651–1674. https://doi.org/10.1177/0192513X15576200

Haveman, R., Mullikin, M. (1999) Alternatives to the official poverty measure: perspectives and assessment, in: *Conference on Poverty: Improving the Definition After*.

Haveman, R., Schwabish, J. (1999) Economic growth and poverty: A return to normalcy? *Focus* 20.

Haven, K., Celaya, M.F., Pierson, J., Weisskopf, A.J., MacKinnon, N.J. (2013) Comparison of health confidence in rural, suburban and urban areas in the UK and the USA: a secondary analysis. *British Medical Journal Open* 3, e002640. https://doi.org/10.1136/bmjopen-2013-002640

Hayes, A.F. (2006) A primer on multilevel modeling. Human Communication Research 32, 385–410.

Hayton, J.C., Allen, D.G., Scarpello, V., 2004. Factor Retention Decisions in Exploratory Factor Analysis: a Tutorial on Parallel Analysis. *Organizational Research Methods* 7, 191–205. https://doi.org/10.1177/1094428104263675

Heaton, T.B. (2002) Factors Contributing to Increasing Marital Stability in the United States. *Journal of Family Issues* 23, 392–409. https://doi.org/10.1177/0192513X02023003004

Heinrich, C.J., Lynn, L.E. (2001) Means and ends: A comparative study of empirical methods for investigating governance and performance. J*ournal of Public Administration Research and Theory* 109–138.

Herman, E. (2008) The American Community Survey: An introduction to the basics. *Government Information Quarterly* 25, 504–519. https://doi.org/10.1016/j.giq.2007.08.006

Hernandez, D.J., Charney, E. (1998) *From Generation to Generation: The Health and Well-Being of Children in Immigrant Families.* National Academies Press, Washington D.C.

Hill, M.S. (1985) The changing nature of poverty. *The Annals of the American Academy of Political and Social Science* 479, 31–47.

Hirschfeld, G., Brachel, R. von, Thielsch, M. (2014) Selecting items for Big Five questionnaires: At what sample size do factor loadings stabilize? *Journal of Research in Personality* 53, 54–63. https://doi.org/10.1016/j.jrp.2014.08.003

Hofferth, S.L., Wissoker, D.A. (1992) Price, Quality, and Income in Child Care Choice. *The Journal of Human Resources* 27, 70–111. https://doi.org/10.2307/145913

Hoffman, C., Paradise, J. (2008) Health Insurance and Access to Health Care in the United States. *Annals of the New York Academy of Sciences* 1136, 149–160. https://doi.org/10.1196/annals.1425.007

Hogarty, K.Y., Hines, C.V., Krombrey, J.D., Ferron, J.M., Mumford, K.R. (2005) The Quality of Factor Solutions in Exploratory Factor Analysis: The Influence of Sample Size, Communality, and Overdetermination. *Educational and Psychological Measurement* 65, 202–226. https://doi.org/10.1177/0013164404267287

Hojman, D., Kast, F. (2009) *On the measurement of poverty dynamics*. HKS Faculty Research Working Paper Series RWP09-035, John F. Kennedy School of Government, Harvard University.

Holgado–Tello, F.P., Chacón–Moscoso, S., Barbero–García, I., Vila–Abad, E. (2010) Polychoric versus Pearson correlations in exploratory and confirmatory factor analysis of ordinal variables. *Quality & Quantity* 44, 153–166. https://doi.org/10.1007/s11135-008-9190-y

Holzer, H.J. (1991) The Spatial Mismatch Hypothesis: What Has the Evidence Shown? *Urban Studies* 28, 105–122. https://doi.org/10.1080/00420989120080071

hooks, bell, 1992. *Black looks: race and representation*. South End Press, Boston, MA.

Horwitz, A.V., McLaughlin, J., White, H.R. (1998) How the Negative and Positive Aspects of Partner Relationships Affect the Mental Health of Young Married People. *Journal of Health and Social Behavior* 39, 124–136. https://doi.org/10.2307/2676395

Hox, J. (1998) Multilevel Modeling: When and Why, in: *Classification, Data Analysis, and Data Highways, Studies in Classification, Data Analysis, and Knowledge Organization*. Springer, Berlin, Heidelberg, pp. 147–154. https://doi.org/10.1007/978-3-642-72087-1_17

Hoyle, R. (2000) Confirmatory Factor Analysis, in: Tinsley, H.E.A., Brown, S.D. (Eds.), *Handbook of Applied Multivariate Statistics and Mathematical Modeling*. Academic Press.

Hoyle, R., Duvall, J. (2004) Determining the Number of Factors in Exploratory and Confirmatory Factor Analysis, in: *The SAGE Handbook of Quantitative Methodology for the Social Sciences*. SAGE Publications, Inc., 2455 Teller Road, Thousand Oaks California 91320 United States of America, pp. 302–317.

Huang, F.L. (2018) Multilevel Modeling and Ordinary Least Squares Regression: How Comparable Are They? *The Journal of Experimental Education* 86, 265–281. https://doi.org/10.1080/00220973.2016.1277339

Huang, F.L. (2016) Alternatives to Multilevel Modeling for the Analysis of Clustered Data. T*he Journal of Experimental Education* 84, 175–196. https://doi.org/10.1080/00220973.2014.952397

Hunt, L.M., Megyesi, M.S. (2008) The Ambiguous Meanings of the Racial/Ethnic Categories Routinely used in Human Genetics Research. *Social Science and Medicine* 66, 349–361. https://doi.org/10.1016/j.socscimed.2007.08.034

Hunting, G., Grace, D., Hankivsky, O. (2015) Taking Action on Stigma and Discrimination: An Intersectionality-Informed Model of Social Inclusion and Exclusion. *Intersectionalities: A Global Journal of Social Work Analysis, Research, Polity, and Practice* 4, 101–125.

Hurtado, A. (1989) Relating to Privilege: Seduction and Rejection in the Subordination of White Women and Women of Color. *Signs* 14, 833–855.

Hutcheson, G. (1999) *The Multivariate Social Scientist*. SAGE Publications, Ltd., 6 Bonhill Street, London EC2A 4PU. https://doi.org/10.4135/9780857028075

Hutto, N., Waldfogel, J., Kaushal, N., Garfinkel, I. (2011) Improving the Measurement of Poverty. *Social Service Review* 85, 39–74. https://doi.org/10.1086/659129

Immergluck, D. (1998) Job proximity and the urban employment problem: Do suitable nearby jobs improve neighbourhood employment rates? *Urban Studies* 35, 7–23.

Islam, M.R. (2015) Rights of the People with Disabilities and Social Exclusion in Malaysia. International *Journal of Social Science and Humanity* 5, 299–305. https://doi.org/10.7763/IJSSH.2015.V5.470

Jaccard, J., Turrisi, R. (2003) Interaction Effects in Multiple Regression, in: *Quantitative Applications in the Social Sciences: Interaction Effects in Multiple Regression*. SAGE Publications, Inc., Thousand Oaks, California. https://doi.org/10.4135/9781412984522

Jackson, C. (1999) Social Exclusion and Gender: Does One Size Fit All? *European Journal of Development Research* 11, 125.

Jackson, C. (1998) Women and Poverty or Gender and Well-Being? *Journal of International Affairs* 52, 67–81.

Jacobs, E. (2007) The Politics of Economic Insecurity (No. 10), *Issues in Governmence Studies*. Brookings Institution.

Johnson, C.K., Mason, P.L. (2012) Theories of Poverty: Traditional Explanations and New Directions, in: Jefferson, P.N. (Ed.), *The Oxford Handbook of the Economics of Poverty*. Oxford University Press, Oxford, pp. 105–136.

Johnson, J.A., Honnold, J.A., Threlfall, P. (2011) Impact of social capital on employment and marriage among low income single mothers. *Journal of Sociology and Social Welfare* 38, 9.

Johnson, K.R., Loscocco, K. (2015) Black Marriage Through the Prism of Gender, Race, and Class. *Journal of Black Studies* 46, 142–171. https://doi.org/10.1177/0021934714562644

Johnson, N.J., Backlund, E., Sorlie, P.D., Loveless, C.A. (2000) Marital Status and Mortality: The National Longitudinal Mortality Study. *Annals of Epidemiology* 10, 224–238. https://doi.org/10.1016/S1047-2797(99)00052-6

Johnson, R.B., Christensen, L. (2013) *Educational Research: Quantitative, Qualitative, and Mixed Approaches.* SAGE Publications.

Juhn, C., Potter, S. (2006) Changes in Labor Force Participation in the United States. *Journal of Economic Perspectives* 20, 27–46. https://doi.org/10.1257/jep.20.3.27

Jung, S. (2013) Exploratory factor analysis with small sample sizes: A comparison of three approaches. *Behavioural Processes* 97, 90–95. https://doi.org/10.1016/j.beproc.2012.11.016

Jung, S., Lee, S. (2011) Exploratory factor analysis for small samples. *Behavior Research Methods* 43, 701–709. https://doi.org/10.3758/s13428-011-0077-9

Kabeer, N. (2005) *Social exclusion: concepts, findings and implications for the MDGs*. Paper commissioned as background for the Social Exclusion Policy Paper, Department for International Development (DFID), London.

Kabeer, N. (2000) Social exclusion, poverty and discrimination towards an analytical framework. *IDS bulletin* 31, 83–97.

Kabeer, N. (1999) Resources, Agency, Achievements: Reflections on the Measurement of Women's Empowerment. *Development and Change* 30, 435–464. https://doi.org/10.1111/1467-7660.00125

Kain, J.F. (1968) Housing Segregation, Negro Employment, and Metropolitan Decentralization. T*he Quarterly Journal of Economics* 82, 175. https://doi.org/10.2307/1885893

Kaiser, H.F., Rice, J. (1974) Little jiffy, mark IV. *Educational and psychological measurement* 34, 111–117.

Kalleberg, A.L. (2012) Job Quality and Precarious Work: Clarifications, Controversies, and Challenges. *Work and Occupations* 39, 427–448. https://doi.org/10.1177/0730888412460533

Kalleberg, A.L., Reskin, B.F., Hudson, K. (2000) Bad Jobs in America: Standard and Nonstandard Employment Relations and Job Quality in the United States. *American Sociological Review* 65, 256–278. https://doi.org/10.2307/2657440

Kaplan, R.M., Chambers, D.A., Glasgow, R.E. (2014) Big Data and Large Sample Size: A Cautionary Note on the Potential for Bias. *Clinical and Translational Science* 7, 342–346. https://doi.org/10.1111/cts.12178

Kearney, M., Harris, B., Jácome, E., Parker, L. (2013) *A Dozen Facts about America's Struggling Lower-Middle Class.* Policy Memo, The Hamilton Project, Brookings Institution, Washington, DC.

Kenney, C. (2004) Cohabiting Couple, Filing Jointly? Resource Pooling and U.S. Poverty Policies. *Family Relations* 53, 11.

Kerner, I. (2012) Questions of intersectionality: Reflections on the current debate in German gender studies. *European Journal of Women's Studies* 19, 203–218. https://doi.org/10.1177/1350506811434666

Kiernan, K. (2002) Cohabitation in Western Europe: trends, issues and implications, in: Booth, A., Crouter, A.C. (Eds.), *Just Living Together: Implications of Cohabitation on Families, Children and Social Policy*. Lawrence Erlbaum Associates Inc, pp. 3–32.

Kim, K., Lee, Yongwoo, Lee, Yu-jeong (2010) A Multilevel Analysis of Factors Related to Poverty in Welfare States. *Social Indicators Research* 99, 391–404. https://doi.org/10.1007/s11205-010-9592-9

Kim, W.H., Lee, J., Bezyak, J. (2017) Five-Year Quality of Life Trajectories for Persons with Disabilities. *Rehabilitation Counseling Bulletin* 61, 8–19. https://doi.org/10.1177/0034355216662201

Kish, L. (1998) Space/time variations and rolling samples. *Journal of Official Statistics* 14, 31–46.

Kish, L. (1990) Rolling Samples and Censuses. *Survey Methodology* 16, 63–79.

Kitano, H.H.L., Yeung, W.-T., Chai, L., Hatanaka, H. (1984) Asian-American Interracial Marriage. *Journal of Marriage and Family* 46, 179–190. https://doi.org/10.2307/351876

Kleiner, M.M., Ham, H. (2002) *Do Industrial Relations Institutions Impact Economic Outcomes?: International and U.S. State-Level Evidence* (NBER Working Paper No. 8729). National Bureau of Economic Research, Inc.

Korbin, J.E. (1992). Child poverty in the United States: introduction. *American Behavioral Scientist* 35, 213–219.

Korenman, S., Miller, J.E., Sjaastad, J.E. (1995) Long-term poverty and child development in the United States: Results from the NLSY. *Children and Youth Services Review* 17, 127–155. https://doi.org/10.1016/0190-7409(95)00006-X

Krahn, G.L., Walker, D.K., Correa-De-Araujo, R. (2015) Persons with Disabilities as an Unrecognized Health Disparity Population. *American Journal of Public Health* 105, S198–S206. https://doi.org/10.2105/AJPH.2014.302182

Kreft, I., de Leeuw, J. (1998) *Introducing multilevel modelling*. Sage, Thousand Oaks, California.

Krivo, L.J., Peterson, R.D., Kuhl, D.C. (2009) Segregation, Racial Structure, and Neighborhood Violent Crime. *American Journal of Sociology* 114, 1765–1802. https://doi.org/10.1086/597285

Kronauer, M. (1998) "Social exclusion" and "underclass" - new concepts for the analysis of poverty, in: Andress, H. (Ed.), E*mpirical Poverty Research in Comparative Perspecitve*. Ashgate, Aldershot, United Kingdom, pp. 51–73.

Kwon, O., Lee, N., Shin, B. (2014) Data quality management, data usage experience and acquisition intention of big data analytics. *International Journal of Information Management* 34, 387–394. https://doi.org/10.1016/j.ijinfomgt.2014.02.002

Labonté, R.N., Hadi, A., Kauffmann, X.E. (2011) *Indicators of Social Exclusion and Inclusion: A Critical and Comparative Analysis of the Literature*. Population Health Improvement Research Network.

Laderchi, C.R., Saith, R., Stewart, F. (2003) Does it matter that we do not agree on the definition of poverty? A comparison of four approaches. *Oxford Development Studies* 31, 243–274.

Lai, M.H.C., Kwok, O. (2015) Examining the Rule of Thumb of Not Using Multilevel Modeling: The "Design Effect Smaller Than Two" Rule. *The Journal of Experimental Education* 83, 423–438. https://doi.org/10.1080/00220973.2014.907229

Laird, J., Parolin, Z., Waldfogel, J., Wimer, C. (2018) Poor State, Rich State: Understanding the Variability of Poverty Rates across U.S. States. *Sociological Science* 5, 628–652. https://doi.org/10.15195/v5.a26

Lamont, M., Lareau, A. (1988) Cultural Capital: Allusions, Gaps and Glissandos in Recent Theoretical Developments. *Sociological Theory* 6, 153–168. https://doi.org/10.2307/202113

Laney, D. (2001) 3-D data management: Controlling data volume, velocity and variety. Application Delivery Strategies by META Group Inc (No. 949). Stamford, CT.

Langer, A., Brown, G.K. (2007) *Cultural status inequalities: An important dimension of group mobilization* (No. Working Paper No. 41). Centre for Research on Inequality, Human Security and Ethnicity, University of Oxford.

Leach, L.S., Butterworth, P., Strazdins, L., Rodgers, B., Broom, D.H., Olesen, S.C. (2010) The limitations of employment as a tool for social inclusion. *BMC Public Health* 10, 621. https://doi.org/10.1186/1471-2458-10-621

LeBlanc, A.J., Tonner, M.C., Harrington, C. (2001) State Medicaid Programs Offering Personal Care Services. *Health Care Financing Review* 22, 155–173.

Ledesma, R.D., Valero-Mora, P. (2007) Determining the number of factors to retain in EFA: An easy-to-use computer program for carrying out parallel analysis. *Practical Assessment, Research & Evaluation* 12, 1–11.

Lee, E.O.J., Brotman, S., 2013. SPEAK OUT! Structural Intersectionality and Anti-Oppressive Practice with LGBTQ Refugees in Canada. *Canadian Social Work Review / Revue Canadienne de Service Social* 30, 157–183.

Lee, J., Cagle, J.G. (2018) Social exclusion factors influencing life satisfaction among older adults. The *Journal of Poverty and Social Justice*; Bristol 26, 35–50. http://dx.doi.org.ezproxy.is.ed.ac.uk/10.1332/175982717X15127351091521

Lee, S.-Y., Song, X.-Y. (2010) Structural Equation Models, in: Peterson, P., Baker, E., McGaw, B. (Eds.), I*nternational Encyclopedia of Education* (Third Edition). Elsevier, Oxford, pp. 453–458. https://doi.org/10.1016/B978-0-08-044894-7.01370-1

Lee, Y.G., Teng, H.-M., Lim, S.-H., Gallo, W.T. (2005) Older Workers: Who are the working poor in the U.S.? *Hallym International Journal of Aging* 7, 95–113.

Lehrer, E.L. (2008) Age at Marriage and Marital Instability: Revisiting the Becker–Landes–Michael Hypothesis. *Journal of Population Economics* 21, 463–484.

Leney, T. (2013) European Approaches to Social Exclusion, in: Hayton, A., Hodgson, A. (Eds.), *Tackling Disaffection and Social Exclusion*. Routledge.

Lenoir, R. (1974) *Les Exclus: Un français sur dix*. Paris: Editions Du Seuil.

Leung, C.W., Ding, E.L., Catalano, P.J., Villamor, E., Rimm, E.B., Willett, W.C. (2012) Dietary intake and dietary quality of low-income adults in the Supplemental Nutrition Assistance Program. *American Journal of Clinical Nutrition* 96, 977–988. https://doi.org/10.3945/ajcn.112.040014

Levinger, G. (1976) A Social Psychological Perspective on Marital Dissolution. *Journal of Social Issues* 32, 21–47. https://doi.org/10.1111/j.1540-4560.1976.tb02478.x

Levitas, R. (2006) The concept and measurement of social exclusion, in: Pantazis, C., Gordon, David., Levitas, R. (Eds.), *Poverty and Social Exclusion in Britain*. The Policy Press, Bristol, pp. 123–160.

Levitas, R., (2005) *The Inclusive Society?: Social Exclusion and New Labour*. Palgrave Macmillan Limited, London, United Kingdom.

Levitas, R. (2004) Let's hear it for Humpty: social exclusion, the third way and cultural capital. *Cultural Trends* 13, 41–56. https://doi.org/10.1080/0954896042000267143

Levitas, R., Pantazis, C., Fahmy, E., Gordon, D., Llloyd, E., Patsios, D. (2007) *The Multidimensional Analysis of Social Exclusion*. Bristol Institute for Public Affairs, University of Bristol. Retrieved from http://dera.ioe.ac.uk/6853/1/multidimensional.pdf.

Leyshon, A., Thrift, N. (1995) Geographies of Financial Exclusion: Financial Abandonment in Britain and the United States. *Transactions of the Institute of British Geographers* 20, 312–341. https://doi.org/10.2307/622654

Lichter, D.T. (2001) *Marriage as public policy*. Progressive Policy Institute, Washington, DC.

Lichter, D.T., Qian, Z., Crowley, M.L. (2005) Child Poverty Among Racial Minorities and Immigrants: Explaining Trends and Differentials. *Social Science Quarterly* 86, 1037–1059. https://doi.org/10.1111/j.0038-4941.2005.00335.x

Lichter, D.T., Qian, Z., Mellott, L.M. (2006) Marriage or Dissolution? Union Transitions Among Poor Cohabiting Women. *Demography* 43, 223–240. https://doi.org/10.1353/dem.2006.0016

Lilja, N., Watts, J. (2010) Rethinking impact: understanding the complexity of poverty and change — overview. *Development in Practice* 20, 917–932.

Lin, A.C., Harris, D.R. (2008) *The Colors of Poverty: Why Racial and Ethnic Disparities Persist*. Russell Sage Foundation New York.

Lin, N. (1999) Building a network theory of social capital. *Connections* 22, 28–51.

Lister, R. (2004) *Poverty*. Blackwell/Polity Press, Cambridge.

Liu, Y., Njai, R.S., Greenlund, K.J., Chapman, D.P., Croft, J.B. (2014) Relationships Between Housing and Food Insecurity, Frequent Mental Distress, and Insufficient Sleep Among Adults in 12 US States, 2009. *Preventing Chronic Disease* 11. https://doi.org/10.5888/pcd11.130334

Lochner, K., Pamuk, E., Makuc, D., Kennedy, B.P., Kawachi, I. (2001). State-level income inequality and individual mortality risk: a prospective, multilevel study. *American Journal of Public Health* 91, 385–391.

Logan, J.R., Oh, S., Darrah, J. (2012) The Political and Community Context of Immigrant Naturalisation in the United States. *Journal of Ethnic and Migration Studies* 38, 535–554. https://doi.org/10.1080/1369183X.2012.659116

Lumley, T., Diehr, P., Emerson, S., Chen, L. (2002) The Importance of the Normality Assumption in Large Public Health Data Sets. *Annual Review of Public Health* 23, 151–169. https://doi.org/10.1146/annurev.publhealth.23.100901.140546

Lutz, W., Goujon, A., K.C., S., Sanderson, W. (2007) Reconstruction of populations by age, sex and level of educational attainment for 120 countries for 1970-2000. *Vienna Yearbook of Population Research* 2007, 193–235. https://doi.org/10.1553/populationyearbook2007s193

Maas, C.J.M., Hox, J.J. (2005) Sufficient Sample Sizes for Multilevel Modeling. *Methodology* 1, 86–92.

MacCallum, R.C., Widaman, K.F., Preacher, K.J., Hong, S. (2001) Sample Size in Factor Analysis: The Role of Model Error. *Multivariate Behavioral Research* 36, 611–637. https://doi.org/10.1207/S15327906MBR3604_06

MacCallum, R.C., Widaman, K.F., Zhang, S., Hong, S. (1999) Sample size in factor analysis. *Psychological Methods* 4, 84.

MacDonald, H. (2006) The American Community Survey: Warmer (more current), but fuzzier (less precise) than the decennial census. *Journal of the American Planning Association* 72, 491–503.

Madden, D. (2000) Relative or Absolute Poverty Lines: A New Approach. *Review of Income & Wealth* 46, 181–199.

Main, G., Bradshaw, J. (2014) *Child poverty and social exclusion: Final report of 2012 PSE study*. University of York, Social Policy Research Unit.

Manovich, L. (2012) Trending: The Promises and the Challenges of Big Social Data, in: Gold, M.K. (Ed.), *Debates in the Digital Humanities*. University of Minnesota Press.

Mansfield, E.R., Helms, B.P. (1982) Detecting Multicollinearity. T*he American Statistician* 36, 158–160. https://doi.org/10.2307/2683167

Martin, S.P. (2006) Trends in Marital Dissolution by Women's Education in the United States. *Demographic Research* 15, 537–560. https://doi.org/10.4054/DemRes.2006.15.20

Massey, D.S. (1996) The age of extremes: Concentrated affluence and poverty in the twenty-first century. *Demography* 33, 395–412. https://doi.org/10.2307/2061773

Massey, D.S., Denton, N.A. (1993) *American apartheid: segregation and the making of the underclass*. Cambridge, Mass.; London: Harvard University Press.

Massey, D.S., Fong, E. (1990) Segregation and Neighborhood Quality: Blacks, Hispanics, and Asians in the San Francisco Metropolitan Area. *Social Forces* 69, 15–32. https://doi.org/10.1093/sf/69.1.15

Mathieson, J., Popay, J., Enoch, E., Escorel, S., Hernandez, M., Johnston, H., Rispel, L. (2008) *Social Exclusion Meaning, measurement and experience and links to health inequalities. A review of literature*. WHO Social Exclusion Knowledge Network Background Paper 1.

Matsunaga, M. (2010) How to factor-analyze your data right: Do's, don'ts, and how-to's. *International Journal of Psychological Research* 3, 97–110.

Matthews, P., Besemer, K. (2014) *Poverty and social networks evidence review*. Joseph Rowntree Foundation, United Kingdom.

McCall, L. (2005) The Complexity of Intersectionality. *Signs: Journal of Women in Culture and Society* 30, 1771–1800. https://doi.org/10.1086/426800

McCarty, A.T. (2016) Child Poverty in the United States: A Tale of Devastation and the Promise of Hope. *Sociol Compass* 10, 623–639. https://doi.org/10.1111/soc4.12386

McKeever, M., Wolfinger, N.H. (2011) Thanks for nothing: Income and labor force participation for never-married mothers since 1982. *Social Science Research* 40, 63–76. https://doi.org/10.1016/j.ssresearch.2010.06.008

McManus, P.A., DiPrete, T.A. (2001) Losers and Winners: The Financial Consequences of Separation and Divorce for Men. *American Sociological Review* 66, 246–268. https://doi.org/10.2307/2657417

McManus, W.S. (1990) Labor Market Effects of Language Enclaves: Hispanic Men in the United States. *The Journal of Human Resources* 25, 228–252. https://doi.org/10.2307/145755

McNeish, D. (2017) Small Sample Methods for Multilevel Modeling: A Colloquial Elucidation of REML and the Kenward-Roger Correction. *Multivariate Behavioral Research*. https://doi.org/10.1080/00273171.2017.1344538

Michels, H.-P. (2013) The "Underclass" Debate – A Discourse that Maligns People Living in Poverty. *Social Change Review* 11, 45–57. https://doi.org/10.2478/scr-2013-0004

Millar, J. (2003) Gender, Poverty and Social Exclusion *Social Policy and Society* 2, 181–188. https://doi.org/10.1017/S1474746403001246

Min, P.G., Kim, C. (2009) Patterns of Intermarriages and Cross-Generational In-Marriages among Native-Born Asian Americans. *International Migration Review* 43, 447–470. https://doi.org/10.1111/j.1747-7379.2009.00773.x

Minkler, M., Stone, R. (1985) The Feminization of Poverty and Older Women. *Gerontologist* 25, 351–357. https://doi.org/10.1093/geront/25.4.351

Moncrieffe, J.M. (2004) *Power Relations, Inequality, and Poverty* (A Concept Paper for Empowerment Team, Poverty Reduction Group World Bank). Overseas Development Institute, London.

Monin, J.K., Clark, M.S. (2011) Why Do Men Benefit More from Marriage Than Do Women? Thinking More Broadly About Interpersonal Processes That Occur Within and Outside of Marriage. *Sex Roles* 65, 320–326. https://doi.org/10.1007/s11199-011-0008-3

Montgomery, D.C., Peck, E.A., Vining, G.G. (2012) *Introduction to Linear Regression Analysis*. John Wiley & Sons Hoboken, New Jersey

Mooney, J.D. (1969) Housing Segregation, Negro Employment and Metropolitan Decentralization: An Alternative Perspective. *The Quarterly Journal of Economics* 83, 299. https://doi.org/10.2307/1883085

Mooney, J.D. (1967) Urban Poverty and Labor Force Participation. *The American Economic Review* 57, 104–119.

Morgan, C., Burns, T., Fitzpatrick, R., Pinfold, V., Priebe, S. (2007) Social exclusion and mental health Conceptual and methodological review. *The British Journal of Psychiatry* 191, 477–483.

Morris, L.A., Deprez, L.S. (2014) The faltering safety net in a reluctant nation: Women's economic security at risk in America. *Women's Studies International Forum, Special Issue on Gender, Mobility and Social Change* 47, 255–268. https://doi.org/10.1016/j.wsif.2013.11.002

Morrison, P.S. (2005) Unemployment and Urban Labour Markets. *Urban Studies* 42, 2261–2288.

Mosse, D. (2007) *Power and the durability of poverty: a critical exploration of the links between culture, marginality and chronic poverty.* Chronic Poverty Research Centre, Manchester.

Mueller, R.O. (1996) Confirmatory Factor Analysis, in: *Basic Principles of Structural Equation Modeling*, Springer Texts in Statistics. Springer, New York, NY, pp. 62–128. https://doi.org/10.1007/978-1-4612-3974-1_2

Mughan, A. (2007) Economic insecurity and welfare preferences: A micro-level analysis. *Comparative Politics* 293–310.

Murray, C. (1996) The Emerging British Underclass, in: Lister, R. (Ed.), *Charles Murray and the Underclass: The Developing Debate*. Civitas: Institute for the Study of Civil Society, London, pp. 23–52.

Murray, C.A. (1994) *Underclass: The Crisis Deepens*. Institute of Economic Affairs, Health and Welfare Unit, London.

Myles, J., Picot, G. (2000) Poverty Indices and Policy Analysis. *Review of Income and Wealth* 46, 161–179. https://doi.org/10.1111/%28ISSN%291475-4991/issues

Nelson, C., Short, K. (2003) *The Distributional Implications of Geographic Adjustment of Poverty Thresholds.* United States Census Bureau. http://www. census. gov/hhes/www/povmeas/papers/geopaper. pdf.

Nezlek, J.B. (2012) Multilevel modeling for psychologists., in: Cooper, H., Camic, P.M., Long, D.L., Panter, A.T., Rindskopf, D., Sher, K.J. (Eds.), A*PA Handbook of Research Methods in*

*Psychology*, Vol 3: Data Analysis and Research Publication. American Psychological Association, Washington, pp. 219–241. https://doi.org/10.1037/13621-011

Nieminen, P., Lehtiniemi, H., Vähäkangas, K., Huusko, A., Rautio, A. (2013) Standardised regression coefficient as an effect size index in summarising findings in epidemiological studies. Epidemiology, *Biostatistics and Public Health* 10. https://doi.org/10.2427/8854

Norris, A.N., Zajicek, A., Murphy-Erby, Y. (2010) Intersectional Perspective and Rural Poverty Research: Benefits, Challenges and Policy Implications. *Journal of Poverty* 14, 55–75. https://doi.org/10.1080/10875540903489413

Notten, G., Neubourg, C. de (2007) *Poverty in Europe and the USA: Exchanging official measurement methods*. Working Paper MGSoG/2007/WP005, Maastricht University

Núñez, A.-M. (2014) Employing Multilevel Intersectionality in Educational Research: Latino Identities, Contexts, and College Access. *Educational Researcher* 43, 85–92. https://doi.org/10.3102/0013189X14522320

OECD (2000) *The Public Employment Service in the United States*. OECD Publishing.

Oroyemi, P., Damioli, G., Barnes, M., Crosier, T. (2010) Understanding the risk of social exclusion across the life course: *Families with children*. A Research Report for the Social Exclusion Task Force, Cabinet Office. Retried from http://natcen.ac.uk/media/25844/understanding-risks-social-exclusion.pdf.

Orshansky, M. (1968) The Shape of Poverty in 1966. *Social Security Bulletin* 44, 219–231.

Orshansky, M. (1965a) Who's Who Among the Poor: A Demographic View of Poverty. *Social Security Bulletin* 28, 3–32.

Orshansky, M. (1965b) Counting the Poor: Another Look at the Poverty Profile. *Social Security Bulletin*; Washington, D.C. 28, 3–29.

Orshansky, M. (1963) Children of the Poor. *Social Security Bulletin* 26, 3–13.

Osberg, L. (2015) How Should One Measure Economic Insecurity? *OECD Statistics Working Papers* No. 2015/01. https://doi.org/10.1787/5js4t78q9lq7-en

Osberg, L., Sharpe, A. (2014) Measuring Economic Insecurity in Rich and Poor Nations. *Review of Income and Wealth* 60, S53–S76. https://doi.org/10.1111/roiw.12114

Osypuk, T.L., Kawachi, I., Subramanian, S.V., Acevedo-Garcia, D. (2006) Are State Patterns of Smoking Different for Different Racial/Ethnic Groups? An Application of Multilevel Analysis. *Public Health Reports* 121, 563–577.

Pain, R. (2000) Place, social relations and the fear of crime: a review. *Progress in Human Geography* 24, 365–387. https://doi.org/10.1191/030913200701540474

Park, M.J., Paul Mulye, T., Adams, S.H., Brindis, C.D., Irwin, C.E. (2006) The Health Status of Young Adults in the United States. *Journal of Adolescent Health* 39, 305–317. https://doi.org/10.1016/j.jadohealth.2006.04.017

Patil, V.H., Singh, S.N., Mishra, S., Todd Donavan, D. (2008) Efficient theory development and factor retention criteria: Abandon the 'eigenvalue greater than one' criterion. *Journal of Business Research* 61, 162–170. https://doi.org/10.1016/j.jbusres.2007.05.008

Patsios, D. (2017) Improvement for some: poverty and social exclusion among older people and pensioners, in: Dermott, E., Main, G. (Eds.), *Poverty and Social Exclusion in the UK: Vol 1: The Nature and Extent of the Problem*. Policy Press.

Paugam, S. (1996) Poverty and Social Disqualification: a Comparative Analysis of Cumulative Social Disadvantage in Europe. *Journal of European Social Policy* 6, 287–303. https://doi.org/10.1177/095892879600600402

Pearce, D. (1993) Change in the other America: Women's poverty in the 1990s. *Women: A Cultural Review* 4, 1–7. https://doi.org/10.1080/09574049308578140

Pearce, D. (1978) The Feminization of Poverty: Women, Work, and Welfare. *Urban and Social Change Review.*

Peng, H., Lu, Y. (2012) Model selection in linear mixed effect models. *Journal of Multivariate Analysis* 109, 109–129. https://doi.org/10.1016/j.jmva.2012.02.005

Peterson, J. (1987) The Feminization of Poverty. *Journal of Economic Issues* 21, 329–337.

Peterson, R.D., Krivo, L.J. (1999) Racial Segregation, the Concentration of Disadvantage, and Black and White Homicide Victimization. *Sociological Forum* 14, 465–493.

Peugh, J.L. (2010) A practical guide to multilevel modeling. Journal of School Psychology 48, 85–112. https://doi.org/10.1016/j.jsp.2009.09.002

Pike, G.R., Rocconi, L.M. (2012) Multilevel modeling: Presenting and publishing the results for internal and external constituents. *New Directions for Institutional Research* 2012, 111–124. https://doi.org/10.1002/ir.20017

Pirani, E. (2011) Evaluating contemporary social exclusion in Europe: a hierarchical latent class approach. *Quality & Quantity* 47, 923–941. https://doi.org/10.1007/s11135-011-9574-2

Pirouz, D.M. (2006) An Overview of Partial Least Squares *SSRN Scholarly Paper* No. ID 1631359. Social Science Research Network, Rochester, NY.

Plasman, R., Rycx, F. (2001) Collective Bargaining and Poverty: A Cross-National Perspective. European *Journal of Industrial Relations* 7, 175–202. https://doi.org/10.1177/095968010172005

Ploeg, M.V., Citro, C.F. (2008) Poverty Measurement: Orshansky's Original Measures and the Development of Alternatives. *Review of Agricultural Economics* 30, 581–590.

Popay, J., Escorel, S., Hernández, M., Johnston, H., Mathieson, J., Rispel, L. (2008) *Defining and measuring social exclusion* Final report to WHO Commission on Social Determinants of Health from Social Exclusion Knowledge Network No. 2.

Prather, C., Fuller, T.R., Jeffries, W.L., Marshall, K.J., Howell, A.V., Belyue-Umole, A., King, W. (2018) Racism, African American Women, and Their Sexual and Reproductive Health: A Review of Historical and Contemporary Evidence and Implications for Health Equity. *Health Equity* 2, 249–259. https://doi.org/10.1089/heq.2017.0045

Proctor, B.D., Semega, J.L., Kollar, M.A. (2016) *Income and poverty in the United States: 2015* Current Population Reports No. P60- 2569. U.S. Census Bureau, Washington D.C.

Prokos, A.H., Keene, J.R. (2010) Poverty Among Cohabiting Gay and Lesbian, and Married and Cohabiting Heterosexual Families. *Journal of Family Issues* 31, 934–959. https://doi.org/10.1177/0192513X09360176

Purdie-Vaughns, V., Eibach, R.P. (2008) Intersectional Invisibility: The Distinctive Advantages and Disadvantages of Multiple Subordinate-Group Identities. *Sex Roles* 59, 377–391. https://doi.org/10.1007/s11199-008-9424-4

Putnam, R.D. (1995) Bowling Alone: America's Declining Social Capital. *Journal of Democracy* 6, 65–78. https://doi.org/10.1353/jod.1995.0002

Qian, Z. (2005) Breaking the Last Taboo: Interracial Marriage in America. *Contexts* 4, 33–37. https://doi.org/10.1525/ctx.2005.4.4.33

Quadagno, J.S. (1994) *The colour of welfare: how racism undermined the war on poverty*. New York ; Oxford : Oxford University Press, 1994.

Rank, M.R. (2006) Toward a New Understanding of American Poverty. *Washington University Journal of Law & Policy* 20, 17.

Rank, M.R., Hirschl, T.A. (2001) Rags or Riches? Estimating the Probabilities of Poverty and Affluence across the Adult American Life Span. *Social Science Quarterly* 82, 651–669.

Rank, M.R., Yoon, H.-S., Hirschl, T.A. (2003) American poverty as a structural failing: Evidence and arguments. *Journal of Sociology & Social Welfare* 30, 3.

Rao, A.R., Monroe, K.B. (1989) The Effect of Price, Brand Name, and Store Name on Buyers' Perceptions of Product Quality: An Integrative Review. *Journal of Marketing Research* 26, 351. https://doi.org/10.2307/3172907

Raphael, S., Smolensky, E. (2009) Immigration and Poverty in The United States. *American Economic Review* 99, 41–44. https://doi.org/10.1257/aer.99.2.41

Ratcliffe, C., McKernan, S.-M., Zhang, S. (2011) How Much Does the Supplemental Nutrition Assistance Program Reduce Food Insecurity? *American Journal of Agricultural Economics* 93, 1082–1098. https://doi.org/10.1093/ajae/aar026

Ravallion, M. (1996) *Issues in measuring and modeling poverty* Policy Research Working Paper Series, World Bank Washington DC. Retrieved from: https://econpapers.repec.org/paper/wbkwbrwps/1615.htm

Razavi, S. (1998) Gendered Poverty and Social Change: An Issues Paper *UNRISD Discussion Paper* No. 94. United Nations Research Institute for Social Development, Switzerland.

Rector, R. (2010) Marriage: America's Greatest Weapon Against Child Poverty. *Backgrounder*. 2465, 1-16

Reeves, M.M. (1990) The States as Polities: Reformed, Reinvigorated, Resourceful. *Annals of the American Academy of Political and Social Science* 509, 83–93.

Reio Jr, T.G., Shuck, B. (2015) Exploratory factor analysis: Implications for theory, research, and practice. *Advances in Developing Human Resources* 17, 12–25.

Reise, S.P., Waller, N.G., Comrey, A.L. (2000) Factor analysis and scale revision. *Psychological Assessment* 12, 287–297. https://doi.org/10.1037//1040-3590.12.3.287

Rejda, G., Haley, J. (2004) Construction of an economic index to measure the causes of economic insecurity. *Forum for Social Economics* 34, 9–30. https://doi.org/10.1007/BF02747501

Ribar, D.C. (1992) Child Care and the Labor Supply of Married Women: Reduced Form Evidence. *The Journal of Human Resources* 27, 134–165. https://doi.org/10.2307/145915

Ridgeway, C.L., Kricheli-Katz, T. (2013) Intersecting Cultural Beliefs in Social Relations: Gender, Race, and Class Binds and Freedoms. *Gender and Society* 27, 294–318.

Ridic, G., Gleason, S., Ridic, O. (2012) Comparisons of Health Care Systems in the United States, Germany and Canada. *Materia Socio-Medica* 24, 112–120. https://doi.org/10.5455/msm.2012.24.112-120

Rippin, N. (2011) A Response to the Weaknesses of the Multidimensional Poverty Index (MPI): The Correlation Sensitive Poverty Index (CSPI) (Briefing Paper No. 19/2011). German Development Institute.

Risman, B.J. (2004) Gender as a Social Structure: Theory Wrestling with Activism. *Gender & Society* 18, 429–450. https://doi.org/10.1177/0891243204265349

Robinson, R.V., Garnier, M.A. (1985) Class Reproduction Among Men and Women in France: Reproduction Theory on Its Home Ground. *American Journal of Sociology* 91, 250–280.

Rodriguez, G. (1999) *From Newcomers to New Americans: The Successful Integration of Immigrants into American Society*. National Immigration Forum, Washington, D.C.

Rolfe, M. (2012) V*oter Turnout: A Social Theory of Political Participation (Political Economy of Institutions and Decisions*). Cambridge University Press, Cambridge. https://doi.org/10.1017/CBO9781139058513

Room, G.J. (1999). Social exclusion, solidarity and the challenge of globalization. *International Journal of Social Welfare* 8, 166–174. https://doi.org/10.1111/1468-2397.00080

Room, G.J. (1995) *Beyond the threshold: the measurement and analysis of social exclusion.* Policy Press, Bristol.

Ross, C.E., Willigen, M.V. (1997) Education and the subjective quality of life. J*ournal of Health and Social Behaviour* 38, 275–97.

Royston, J.P. (1983) A Simple Method for Evaluating the Shapiro–Francia W′ Test of Non-Normality. *Journal of the Royal Statistical Society*: Series D (The Statistician) 32, 297–300. https://doi.org/10.2307/2987935

Royston, P. (1992) Approximating the Shapiro-Wilk W-test for non-normality. *Statistics and Computing* 2, 117–119. https://doi.org/10.1007/BF01891203

Ruggles, P. (2008) Poverty Measurement and Politics: Why Is It So Hard to Change the Way We Measure Poverty? *Review of Agricultural Economics* 30, 591–599. https://doi.org/10.1111/j.1467-9353.2008.00433.x

Russom, P. (2011) Big Data Analytics *TDWI Best Practices* No. Fourth quarter. Retrieved from: https://tdwi.org/research/2011/09/best-practices-report-q4-big-data-analytics/asset.aspx?tc=assetpg&m=1

Saatcioglu, B., Corus, C. (2014) Poverty and Intersectionality: A Multidimensional Look into the Lives of the Impoverished. *Journal of Macromarketing* 34, 122–132. https://doi.org/10.1177/0276146713520600

Sagiroglu, S., Sinanc, D. (2013) Big data: A review. Presented at the 2013 International Conference on Collaboration Technologies and Systems (CTS), San Diego, CA, pp. 42–47. https://doi.org/10.1109/CTS.2013.6567202

Saltkjel, T., Dahl, E., Wel, K.A. van der (2013) Health related social exclusion in Europe: a multilevel study of the role of welfare generosity. *International Journal for Equity in Health* 12, 81. https://doi.org/10.1186/1475-9276-12-81

Sampson, R.J. (2003) The Neighborhood Context of Well-Being. *Perspectives in Biology and Medicine* 46, S53–S64. https://doi.org/10.1353/pbm.2003.0073

Sampson, R.J., Lauritsen, J.L. (1997) Racial and Ethnic Disparities in Crime and Criminal Justice in the United States. *Crime and Justice* 21, 311–374.

Sanchez, T.W. (1999) The Connection Between Public Transit and Employment. *Journal of the American Planning Association* 65, 284–296. https://doi.org/10.1080/01944369908976058

Sarstedt, M., Mooi, E. (2014) Factor Analysis, in: *A Concise Guide to Market Research*. Springer Berlin Heidelberg, Berlin, pp. 235–272. https://doi.org/10.1007/978-3-642-53965-7_8

Sassler, S., Miller, A.J., (2011). Class Differences in Cohabitation Processes. *Family Relations* 60, 163–177.

Saunders, P.G. (2013) Reflections on the Concept of Social Exclusion and the Australian Social Inclusion Agenda. *Social Policy & Administration* 47, 692–708. https://doi.org/10.1111/spol.12038

Saunders, P.G. (2003) Can social exclusion provide a new framework for measuring poverty? University of New South Wales, Social Policy Research Centre

Sayer, L.C., England, P., Bittman, M., Bianchi, S.M. (2009) How Long Is the Second (Plus First) Shift? Gender Differences in Paid, Unpaid, and Total Work Time in Australia and the United States. *Journal of Comparative Family Studies*; Calgary 40, 523–IX.

Schalock, R.L. (2004) The concept of quality of life: what we know and do not know. *Journal of Intellectual Disability Research* 48, 203–216. https://doi.org/10.1111/j.1365-2788.2003.00558.x

Schierup, C.-U. (2010)" Diversity" and Social Exclusion in Third Way Sweden: The" Swedish Model" in Transition, 1975 – 2005 (No. 35), *Themes on Migration and Ethnic Studies.* Institute for Research on Migration, Ethnicity and Society, Sweden.

Schill, M.H., Friedman, S., Rosenbaum, E. (1998) The Housing Conditions of Immigrants in New York City. *Journal of Housing Research* 9, 201–231.

Schmeiser, M.D. (2012) The impact of long-term participation in the supplemental nutrition assistance program on child obesity. *Health Economics* 21, 386–404. https://doi.org/10.1002/hec.1714

Schmitt, N. (1996) Uses and Abuses of Coefficient Alpha. *Psychological Assessment* 8, 350–353.

Schmitt, T.A., Sass, D.A. (2011) Rotation Criteria and Hypothesis Testing for Exploratory Factor Analysis: Implications for Factor Pattern Loadings and Interfactor Correlations. *Educational and Psychological Measurement* 71, 95–113. https://doi.org/10.1177/0013164410387348

Schwanen, T., Dijst, M. (2002) Travel-time ratios for visits to the workplace: the relationship between commuting time and work duration. *Transportation Research Part A: Policy and Practice* 36, 573–592. https://doi.org/10.1016/S0965-8564(01)00023-4

Seccombe, K. (2000) Families in Poverty in the 1990s: Trends, Causes, Consequences, and Lessons Learned. *Journal of Marriage and Family* 62, 1094–1113.

Seddon, T. (2006) Drugs, Crime and Social ExclusionSocial Context and Social Theory in British Drugs–Crime Research. *British Journal of Criminology* 46, 680–703. https://doi.org/10.1093/bjc/azi079

Segal, E.A., Kilty, K.M., Kim, R.Y. (2002) Social and Economic Inequality and Asian Americans in the United States. *Journal of Poverty* 6, 5–21. https://doi.org/10.1300/J134v06n04_02

Segal, E.A., Peck, L.R. (2006) Sequential Costs of Poverty: What Traditional Measures Overlook, The Special Issue with Coping with Poverty. *Journal of Sociology & Social Welfare* 33, 227–238.

Sen, A. (2000) Social Exclusion: Concept, Application and Scrutiny. Asian Development Bank, *Social Development Papers* No. 1.

Sen, A. (1999) *Development as Freedom* Oxford: Oxford University Press, 1999.

Shields, S.A. (2008) Gender: An Intersectionality Perspective. *Sex Roles* 59, 301–311. https://doi.org/10.1007/s11199-008-9501-8

Shihadeh, E.S., Flynn, N. (1996) Segregation and Crime: The Effect of Black Social Isolation on the Rates of Black Urban Violence. *Social Forces* 74, 1325–1352. https://doi.org/10.1093/sf/74.4.1325

Sigle-Rushton, W., McLanahan, S. (2002) For richer or poorer? Marriage as an anti-poverty strategy in the United States. *Population* 57, 519. https://doi.org/10.3917/popu.203.0519

Silver, H. (1994) Social exclusion and social solidarity: Three paradigms. *International Labour Review* 133, 531.

Silver, H., Miller, S.M. (2003) Social exclusion. *Indicators* 2, 5–21.

Silver, H.F. (2007) The process of social exclusion: the dynamics of an evolving concept *CPRC Working Paper* No. 95. Chronic Poverty Research Centre, Dept. of Sociology, Brown University, Providence, RI.

Slesnick, D.T. (1991) The Standard of Living in the United States. *Review of Income and Wealth* 37, 363–386. https://doi.org/10.1111/j.1475-4991.1991.tb00379.x

Sloane, P.J., Jones, M.K. (2012) Disability and Social Exclusion, in: Parodi, G., Sciulli, D. (Eds.), *Social Exclusion*. Physica-Verlag HD, Heidelberg, pp. 127–148. https://doi.org/10.1007/978-3-7908-2772-9_7

Smeeding, T., Thévenot, C., 2016. Addressing Child Poverty: How Does the United States Compare With Other Nations? Academic Pediatrics, Child Poverty in the United States 16, S67–S75. https://doi.org/10.1016/j.acap.2016.01.011

Smeeding, T.M. (1977) The Antipoverty Effectiveness of In-Kind Transfers. *The Journal of Human Resources* 12, 360–378. https://doi.org/10.2307/145496

Smeeding, T.M., Saunders, P., Coder, J., Jenkins, S., Fritzell, J., Hagenaars, A.J., Hauser, R., Wolfson, M. (1993) Poverty, Inequality, and Family Living Standards Impacts across Seven Nations: The Effect of Noncash Subsidies for Health, Education and Housing. *Review of Income and Wealth* 39, 229–256.

Smock, P.J., Gupta, S. (2002) Cohabitation in contemporary North America., in: Booth, A., Crouter, A.C. (Eds.), *Just Living Together: Implications of Cohabitation on Families, Children and Social Policy*. Lawrence Erlbaum Associates Inc, pp. 3–32.

Snowden, L.R., Thomas, K. (2000) Medicaid and African American Outpatient Mental Health Treatment. *Mental Health Services Research* 2, 115–120.

Sonowal, C. (2008) Indian Tribes and Issue of Social Inclusion and Exclusion. *Studies of Tribes and Tribals* 6, 123–134. https://doi.org/10.1080/0972639x.2008.11886586.

Sparkes, J. (1999) Schools, education and social exclusion *STICERD Research Paper* No. CASE029). Retrieved from: https://ssrn.com/abstract=1158920.

Stafford, M., Marmot, M. (2003) Neighbourhood deprivation and health: does it affect us all equally? *International Journal of Epidemiology* 32, 357–366. https://doi.org/10.1093/ije/dyg084

Stapleton, D.C., O'Day, B.L., Livermore, G.A., Imparato, A.J. (2006) Dismantling the Poverty Trap: Disability Policy for the Twenty-First Century. *Milbank Quarterly* 84, 701–732. https://doi.org/10.1111/j.1468-0009.2006.00465.x

Starrels, M.E., Bould, S., Nicholas, L.J. (1994) The Feminization of Poverty in the United Sates: Gender, Race, Ethnicity, and Family Factors. *Journal of Family Issues* 15, 590–607.

Stewart, F. (2005) Horizontal inequalities: a neglected dimension of development. In: *Wider Perspectives on Global Development*. Studies in Development Economics and Policy. Palgrave Macmillan, London

Stolzenberg, R.M., Tienda, M. (1997) English Proficiency, Education, and the Conditional Economic Assimilation of Hispanic and Asian Origin Men. *Social Science Research* 26, 25–51. https://doi.org/10.1006/ssre.1996.0588

Subramanian, S.V., Chen, J.T., Rehkopf, D.H., Waterman, P.D., Krieger, N. (2005) Racial Disparities in Context: A Multilevel Analysis of Neighborhood Variations in Poverty and Excess Mortality Among Black Populations in Massachusetts. *American Journal of Public Health* 95, 260–265. https://doi.org/10.2105/AJPH.2003.034132

Sullivan, A. (2001) Cultural Capital and Educational Attainment. *Sociology* 35, 893–912.

Sullivan, D.H., Ziegert, A.L. (2008) Hispanic Immigrant Poverty: Does Ethnic Origin Matter? *Population Research and Policy Review* 27, 667. https://doi.org/10.1007/s11113-008-9096-3

Suro, R., Wilson, J.H., Singer, A. (2011) Immigration and Poverty in America's Suburbs, *Metropolitan Opportunity Series*. Brookings Institute, Washington, D.C.

Sykes, J., Križ, K., Edin, K., Halpern-Meekin, S. (2015) Dignity and Dreams: What the Earned Income Tax Credit (EITC) Means to Low-Income Families. *American Sociological Review* 80, 243–267. https://doi.org/10.1177/0003122414551552

Tavakol, M., Dennick, R. (2011) Making sense of Cronbach's alpha. *International Journal of Medical Education* 2, 53–55. https://doi.org/10.5116/ijme.4dfb.8dfd

TechAmerica Foundation's Federal Big Data Commission (2012) *Demystifying big data: A practical guide to transforming the business of Government.* Retrieved from: https://bigdatawg.nist.gov/_uploadfiles/M0068_v1_3903747095.pdf.

Thiede, B.C., Kim, H., Slack, T. (2017) Marriage, Work, and Racial Inequalities in Poverty: Evidence from the United States. *Family Relations* 79, 1241–1257. https://doi.org/10.1111/jomf.12427

Thomas, S.L., Heck, R.H. (2001). Analysis of Large-Scale Secondary Data in Higher Education Research: Potential Perils Associated with Complex Sampling Designs. *Research in Higher Education* 42, 517–540.

Thomas, S.L., Heck, R.H., Bauer, K.W. (2005) Weighting and adjusting for design effects in secondary data analyses. *New Directions for Institutional Research* 2005, 51–72. https://doi.org/10.1002/ir.155

Thornton, A. (1985). Changing Attitudes Toward Separation and Divorce: Causes and Consequences. *American Journal of Sociology* 90, 856–872.

Throsby, D. (1999) Cultural Capital. *Journal of Cultural Economics* 23, 3–12. https://doi.org/10.1023/A:1007543313370

Tiehen, L., Ver Ploeg, M. (2012) S*NAP Benefits Alleviate the Intensity and Incidence of Poverty*. US Department of Agriculture Economic Research Service, Washington, D.C.

Titheridge, H., Mackett, R.L., Christie, N., Oviedo Hernández, D., Ye, R. (2014) T*ransport and poverty: a review of the evidence.* (UCLTI Publications). UCL Transport Institute, University College London: London, UK

Tomaskovic-Devey, D. (1987) Labor Markets, Industrial Structure, and Poverty: A Theoretical Discussion and Empirical Example. *Rural Sociology* 52, 56–74.

Townsend, P. (2006) What is Poverty? An Historical Perspective. *Poverty in Focus* 5–6.

Townsend, P. (1979) *Poverty in the United Kingdom: A Survey of Household Resources and Standards of Living*, First Edition edition. ed. Penguin Books Ltd, Harmondsworth, Eng.; New York.

Treas, J. (1987) The Effect of Women's Labor Force Participation on the Distribution of Income in the United States. *Annual Review of Sociology* 13, 259–288. https://doi.org/10.1146/annurev.so.13.080187.001355

Treas, J., Mazumdar, S. (2002) Older people in America's immigrant families: Dilemmas of dependence, integration, and isolation. *Journal of Aging Studies* 16, 243–258. https://doi.org/10.1016/S0890-4065(02)00048-8

Triest, R. (1997) Regional Differences in Family Poverty. *New England Economic Review* 3–17.

Turner, T., Niemeier, D. (1997) Travel to work and household responsibility: new evidence. *Transportation* 24, 397–419. https://doi.org/10.1023/A:1004945903696

Umberson, D., Montez, J.K. (2010) Social Relationships and Health: A Flashpoint for Health Policy. *Journal of Health and Social Behaviour* 51, S54–S66. https://doi.org/10.1177/0022146510383501

United Nations (2009) *Overcoming Economic Insecurity*. World Economic and Social Survey 2008 - Academic Foundation.

U.S. Census Bureau (2018) *Understanding and Using American Community Survey Data: What All Data Users Need to Know*. U.S. Census Bureau, Washington D.C.

U.S. Census Bureau (2016a) New Census Data Show Differences Between Urban and Rural Populations (Press Release No. CB16-210). U.S. Census Bureau, Washington, D.C.

U.S. Census Bureau (2016b). *PUMS: Accuracy of the Data 2015*. U.S. Census Bureau, Washington, D.C.

U.S. Census Bureau (2016c) 2015 American Community Survey Public Use Microdata Files, electronic dataset and documentation. American FactFinder, United States Census Bureau, Washington D.C. Available at: https://factfinder.census.gov/faces/nav/jsf/pages/searchresults.xhtml?refresh=t#

U.S. Census Bureau (2014). *American Community Survey Design and Methodology Report*. U.S. Census Bureau, Washington, D.C.

U.S. Census Bureau (2008) *A Compass for Understanding and Using American Community Survey Data*. U.S. Government Printing Office, Washington, D.C.

U.S. Census Bureau (2008) *Income, Earnings, and Poverty Data from The 2007 American Community Survey*. Government Accountability Office, Washington D.C.

U.S. Department of Housing and Urban Development (2007) *Measuring Overcrowding in Housing.* Office of Policy Development and Research Publication, Washington, D.C.

Valdez, C.R., Valentine, J.L., Padilla, B. (2013) "Why We Stay": Immigrants' motivations for remaining in communities impacted by anti-immigration policy. *Cultural Diversity Ethnic Minority Psychology* 19, 279–287. https://doi.org/10.1037/a0033176

Van Auken, P.M., Hammer, R.B., Voss, P.R., Veroff, D.L. (2006) The American Community Survey in counties with "seasonal" populations. *Population Research and Policy Review* 25, 275–292. https://doi.org/10.1007/s11113-006-0010-6

van der Eijk, C., Rose, J. (2015) Risky Business: Factor Analysis of Survey Data – Assessing the Probability of Incorrect Dimensionalisation. *PLoS One* 10. https://doi.org/10.1371/journal.pone.0118900

van Deursen, A.J., van Dijk, J.A. (2014) The digital divide shifts to differences in usage. *New Media & Society* 16, 507–526. https://doi.org/10.1177/1461444813487959

Vijaya, R.M., Lahoti, R., Swaminathan, H. (2014) Moving from the Household to the Individual: Multidimensional Poverty Analysis. *World Development* 59, 70–81. https://doi.org/10.1016/j.worlddev.2014.01.029

Vranken, J. (2001) Unravelling the Social Strands of Poverty, in: Andersen, H.T., Van Kempen, R. (Eds.), *Governing European Societies: Social Fragementation, Social Exclusion, and Urban Governance*. Ashgate Publishing Limited, England, pp. 71–92.

Waddell, G., Burton, A.K. (2007) *Is work good for your health and well-being?* The Stationery Office, London.

Wagle, U.R. (2009) Capability Deprivation and Income Poverty in the United States, 1994 and 2004: Measurement Outcomes and Demographic Profiles. *Social Indicators Research*; Dordrecht 94, 509–533. http://dx.doi.org.ezproxy.is.ed.ac.uk/10.1007/s11205-009-9446-5

Walker, A. (1980) The Social Creation of Poverty and Dependency in Old Age. *Journal of Social Policy* 9, 49–75. https://doi.org/10.1017/S0047279400009545

Wang, X., Feng, H., Xia, Q., Alkire, S. (2016) O*n the relationship between income poverty and multidimensional poverty in China*. Oxford Poverty & Human Development Initiative, Oxford.

Warner, T.D., Manning, W.D., Giordano, P.C., Longmore, M.A. (2011) Relationship formation and stability in emerging adulthood: do sex ratios matter? *Social Forces* 90.

Waters, M.C., Eschbach, K. (1995) Immigration and Ethnic and Racial Inequality in the United States. *Annual Review of Sociology* 21, 419–446. https://doi.org/10.1146/annurev.so.21.080195.002223

Watt, P., Jacobs, K. (2000) Discourses of Social Exclusion: An Analysis of Bringing Britain Together: A National Strategy for Neighbourhood Renewal. *Housing, Theory and Society* 17, 14–26. https://doi.org/10.1080/140360900750044746

Weathers II, R.R., Bruyere, S.M., Houtenville, A.J., Stapleton, D.C. (2005) A Guide to Disability Statistics from the American Community Survey, *Disability Statistics User Guide Series*. Cornell University, Employment and Disability Institute.

Weizmann-Henelius, G., Putkonen, H., Grönroos, M., Lindberg, N., Eronen, M., Häkkänen-Nyholm, H. (2010) Examination of psychopathy in female homicide offenders — Confirmatory factor analysis of the PCL-R. *International Journal of Law and Psychiatry* 33, 177–183. https://doi.org/10.1016/j.ijlp.2010.03.008

Wells, B., Zinn, M.B. (2004) The Benefits of Marriage Reconsidered J*ournal of Sociology & Social Welfare* 31, 59–80.

Western, B., Bloome, D., Sosnaud, B., Tach, L. (2012) Economic Insecurity and Social Stratification. *Annual Review of Sociology* 38, 341–359. https://doi.org/10.1146/annurev-soc-071811-145434

White, L., Rogers, S.J. (2000) Economic Circumstances and Family Outcomes: A Review of the 1990s. *Journal of Marriage and Family* 62, 1035–1051. https://doi.org/10.1111/j.1741-3737.2000.01035.x

White-Means, S.I., Osmani, A.R. (2018) Affordable Care Act and Disparities in Health Services Utilization among Ethnic Minority Breast Cancer Survivors: Evidence from Longitudinal Medical Expenditure Panel Surveys 2008‑2015. *International Journal Environmental Research and Public Health* 15. https://doi.org/10.3390/ijerph15091860

Williams, J.D. (2013) The American Community Survey: Development, implementation, and issues for congress, *CRS Report for Congress* No. 7–5700, Congressional Research Service, Washington D.C.

Willson, A.E., Hardy, M.A. (2002) Racial disparities in income security for a cohort of aging American women. *Social Forces* 80, 1283–1306.

Wilson, W.J. (1987) *The Truly Disadvantaged: The Inner City, the Underclass, and Public Policy*. Chicago ; London : University of Chicago Press, 1987.

Wilson, W.J. (1984) The Black Underclass. *The Wilson Quarterly* 8, 88–99.

Wolf, J.P. (2007) Sociological Theories of Poverty in Urban America. *Journal of Human Behavior in the Social Environment* 16, 41–56. https://doi.org/10.1300/J137v16n01_04

Wood, J.M., Tataryn, D.J., Gorsuch, R.L. (1996) Effects of under-and overextraction on principal axis factor analysis with varimax rotation. *Psychological Methods* 1, 354.

Worts, D., Sacker, A., McDonough, P. (2010) Falling Short of the Promise: Poverty Vulnerability in the United States and Britain, 1993–2003. *American Journal of Sociology* 116, 232–271. https://doi.org/10.1086/653542

Wright, E.O., Rogers, J. (2010) *American Society: How it Really Works*. W.W. Norton & Company.

Wroe, A. (2016) Economic insecurity and political trust in the united states. *American Politics Research* 44, 131–163.

Yang, Y., Xia, Y. (2015) On the number of factors to retain in exploratory factor analysis for ordered categorical data. *Behavioral Research* 47, 756–772. https://doi.org/10.3758/s13428-014-0499-2

Yeo, R., Moore, K. (2003) Including Disabled People in Poverty Reduction Work: "Nothing About Us, Without Us." *World Development* 31, 571–590. https://doi.org/10.1016/S0305-750X(02)00218-8

Yong, A.G., Pearce, S. (2013) A beginner's guide to factor analysis: Focusing on exploratory factor analysis. *Tutorials in Quantitative Methods for Psychology* 9, 79–94.

Yumatle, C. (2015) Pluralism, in: Gibbons, M.T. (Ed.), *The Encyclopedia of Political Thought*. John Wiley & Sons, Ltd, p. 19.

Yuval-Davis, N. (2006) Intersectionality and Feminist Politics. *European Journal of Women's Studies* 13, 193–209. https://doi.org/10.1177/1350506806065752

Zagorsky, J.L. (2005) Marriage and divorce's impact on wealth. *Journal of Sociology* 41, 406–424. https://doi.org/10.1177/1440783305058478

# APPENDIX A

The purpose of this appendix is to acknowledge additional data sources that are available for the American context. Five data sets are presented in Table A1. Most of the data sets discussed in the table are publicly available, though the Panel Study of Income Dynamics has components of it that are restricted and requires a special contract. Each data set includes information about income, education, and employment, but are not as ideal as the ACS. For instance, the Health and the Retirement Survey only samples older adults in the United States. The analysis presented in this thesis focuses on individuals aged 18 and up (for reasons stated in Chapter 5). Additionally, the Survey of Income and Program Participation explores public health mainly, as it seeks to analyse the relationship between diet, nutrition, and public health. None of the datasets presented in Table A1 have the sample size as substantial as the ACS. With these options explored, I am confident of the suitability of the ACS to explore and assess the research questions presented throughout this thesis.

**Table A 1: Additional datasets to measure social exclusion in the United States**

| Data | Relevant B-SEM indicators | Sample | Other notable information |
|---|---|---|---|
| Health and Retirement Survey | Internet use, family structure, disability, health and health insurance, household characteristics, assets | Older adults (ages 55-64) | The data are released after a longer lag time than the ACS data. For instance, data collected in 2014 were not released until 2017. |
| Survey of Income and Program Participation | Participation in government programs, such as Social Security and other income assistance, health insurance and health insurance, household structure, income, assets, disability, education | Individuals 15 and up from sampled households. SIPP sample size ranges between 14,000 and 52,000 households per wave. | The SIPP is typically used to analyse the use and eligibility of government programs and evaluate how modifications of changes to programs impact households. |

| *Part 2*<br><br>**Data** | **Relevant B-SEM indicators** | **Sample** | **Other notable information** |
|---|---|---|---|
| National Survey of Health and Nutrition Survey | Health-related questions, including disability, cognition, smoking use. Additional questions related to food security, health insurance, housing characteristics | All individuals in selected households. The sample represents the total non-institutionalised American population in the 50 US states and the District of Columbia. The sample design is multi-stage like the ACS. | A key goal of this data set is to study the relationship between nutrition, diet, and health. The focus is on public health, which the collected data relate to. |
| Panel Study of Income Dynamics | Consumption and expenditures, income and transfers, education, housing, health status, employment, marriage, health insurance | The sample size began with about 5,000 families (about 18,000 people) in 1968. The sample size has grown as the children of early families created their own families. It is estimated that in 2017, the PSID contains data for 10,000 families (24,000 individuals). | |
| Current Population Survey | Disability status, marital status, employment, weekly and hourly earnings, industry of employment, multiple job status | Approximately 62,000 housing units each month | The March Supplement of the Current Population Survey is used to determine poverty status in the United States. |

# APPENDIX B

I offer this Appendix to highlight that I have considered if the factors of disadvantage presented in the thesis hold in other years. Utilising the same indicators from the B-SEM acknowledged in Chapter 5 for the year 2015, I conducted an exploratory factor analysis using data from the 2008 ACS PUMS files. The results of the EFA are presented in Table B1.

Table B1 shows that five items loaded onto factor one (‘*labour force participation’)*: disability status, travel time to work, not in the labour force, employed civilian, and driving as a means of transportation to work. Six items loaded onto factor 2 (‘*economic security*’): health insurance coverage, overcrowded housing, relative income poverty, educational attainment, total individual income, and being a recipient of food stamps. Three items loaded on the third factor (‘*marriage as a social resource*’): married, divorced and unmarried partners.

### Table B 1: Obliquely Rotated Factor Loadings 2008 EFA

| Factor | Labour Force Participation | Economic Security | Marriage: a Social Resource | Uniqueness |
|---|---|---|---|---|
| Disability | -0.35 | | | 0.85 |
| Health Insurance | | 0.43 | | 0.82 |
| Overcrowded housing | | -0.41 | | 0.86 |
| Relative income poverty | | -0.56 | | 0.61 |
| Travel time to work | 0.53 | | | 0.70 |
| Not in Labour Force | -0.94 | | | 0.13 |
| Employed civilian | 0.95 | | | 0.09 |
| Means of transportation to work: drive | 0.82 | | | 0.31 |
| Educational attainment | | 0.50 | | 0.74 |
| Married | | | -0.60 | 0.60 |
| Divorced | | | 0.55 | 0.72 |
| Unmarried Partners | | | 0.35 | 0.86 |
| Total individual income | | 0.44 | | 0.75 |
| Food stamps | | -0.37 | | 0.85 |
| Variance | 3.13 | 1.74 | 0.97 | |
| Proportion of variance | 0.70 | 0.39 | 0.22 | |

Source: 2008 ACS PUMS

The resulting factors using the 2008 and 2015 ACS PUMS are consistent. They both produce a three-factor solution with similar factor loadings. In Chapter 4, I acknowledged that addresses are not sampled for the ACS for at least five years. Because there is a seven-year gap between 2008 and 2015, it is possible that some addresses utilised the 2015 data were sampled in 2008. However, the results for both EFAs offer some evidence that before the Financial Crisis and after, the dimensions of disadvantage in the United States were relatively similar (for this group of individuals, using the available indicators from the ACS PUMS).